PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 3/00, 13/00, C07H 21/00
C12P 21/06, 21/02, 21/04
C12N 15/00

A1

(11) International Publication Number: **WO 92/14748**

(43) International Publication Date: 3 September 1992 (03.09.92)

(21) International Application Number: PCT/US92/01300

(22) International Filing Date: 20 February 1992 (20.02.92)

(30) Priority data:
657,236 22 February 1991 (22.02.91) US

(60) Parent Application or Grant
(63) Related by Continuation
US 657,236 (CIP)
Filed on 22 February 1991 (22.02.91)

(71) Applicant *(for all designated States except US):* AMERICAN CYANAMID COMPANY [US/US]; 1937 West Main Street, P.O. Box 60, Stamford, CT 06904 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : TERMAN, Bruce, Israel [US/US]; 17 Prospect Street, Monroe, NY 10950 (US). CARRION, Miguel, Eduardo [EC/US]; 26 Summit Avenue, Spring Valley, NY 10977 (US).

(74) Agent: GORDON, Alan, M.; American Cyanamid Company, 1937 West Main Street, P.O. Box 60, Stamford, CT 06904 (US).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, KR, LU (European patent), MC (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.

(54) Title: IDENTIFICATION OF A NOVEL HUMAN RECEPTOR TYROSINE KINASE GENE

(57) Abstract

A DNA sequence encoding a novel human growth factor receptor referred to as a type III receptor tyrosine kinase is described. The amino acid sequence of the receptor is also described. The receptor has a sequence which is similar to that of the kinase domains of known type III receptor tyrosine kinases, but which is unique in its kinase insert domain sequence. The receptor binds specifically to the vascular endothelial cell growth factor.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CS	Czechoslovakia
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
PL	Poland
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

IDENTIFICATION OF A NOVEL HUMAN RECEPTOR TYROSINE KINASE GENE

FIELD OF THE INVENTION

This invention relates to the DNA sequence encoding a novel human growth factor receptor which is a type III receptor tyrosine kinase. The receptor is referred to as Kinase insert Domain containing Receptor (KDR) and binds specifically to the growth factor vascular endothelial cell growth factor (VEGF). This invention also relates to the amino acid sequence of the receptor.

BACKGROUND OF THE INVENTION

Growth factors are small molecules which regulate normal cell growth and development through interaction with cell surface receptors. The receptors for a number of growth factors are referred to as tyrosine kinases; that is, binding of growth factor to the receptor stimulates an increased phosphorylation of tyrosine amino acids within the receptor; this is turn leads to cellular activation (Bibliography 1).

There is increasing evidence that genetic alterations affecting the expr ssion of receptor tyrosine kinases (RTK) can contribute to the altered cell growth associat d with cancer. This conclusion is

supported by the frequent identification of RTK as products of the oncogenes for many of the acutely transforming retroviruses (e.g., 2,3,4) and the overexpression of RTK in certain cancers (5). The identification of a novel RTK may lead to a better understanding of cell growth under both normal and transforming circumstances.

The amino acid sequence in the catalytic domain of all tyrosine kinases has been conserved (6). Detailed analysis of the amino acid sequences within the catalytic and noncatalytic domains of RTK indicates the existence of distinct structural subtypes. One group of RTK (designated type III) includes the ckit proto-oncogene and the receptors for platelet derived growth factor (PDGF) and colony stimulating factor-1 (CSF-1).

The most unusual feature of this subtype is that its catalytic (kinase) domain is interrupted by a long insertion sequence of 12-102 amino acids (the kinase insert domain) The two peptides constituting the kinase domain are conserved between the receptors, while the sequence of the kinase insert domain is unique for each receptor.

Several approaches have been tried in order to identify novel RTK, including low-stringency screening of cDNA libraries with previously characterized DNA probes (7). More recently, a technique has been developed that is capable of greatly facilitating the identification of novel genes for which some sequence data are known. The polymerase chain reaction (PCR) has been used to identify novel members of several gene families including those of guanine nucleotide regulatory prot ins (8) and protein phosphatases (9). PCR has been us d to identify nov l tyrosine kinase gen s (10), though th primers us d in

that study were designed from DNA segments contained in all tyrosine kinases, rather than being specifically directed against RTK. It is a continuing goal to identify receptors for growth factors.

The elucidation of the growth factors, as well as their receptors, involved in regulating endothelial cell function is critical for the understanding of how new blood vessels are formed (angiogenesis). Angiogenesis plays a significant role in both normal and pathological events such as embryogenesis, progression of ocular diseases, and wound healing (11). In particular, angiogenesis is an important process for the growth of tumors (11). Angiogenesis is a complex process involving endothelial cell proliferation, migration, and tissue infiltration. These events are stimulated by growth factors which either (i) act directly on endothelial cells (12,13), or (ii) act indirectly by inducing host cells to release specific endothelial cell growth factors (11). One member of the first group is vascular endothelial cell growth factor (VEGF), also known as vascular permeability factor (14-16). Besides its angiogenic activity, VEGF displays the physiological function of increasing the permeability of capillary vessels to different macromolecules (14).

SUMMARY OF THE INVENTION

The present invention relates to novel DNA segments which together comrpise a gene which encodes type III RTK. The type III RTK encoded by the gene is designated the KDR protein (which stands for Kinase ins rt Domain containing Rec ptor). The KDR prot in binds specifically to th growth factor VEGF (vascular endoth lial cell growth factor).

The DNA segments are identified and isolated through the use of PCR technology. The overall strategy is summarized as follows:

PCR is used to amplify the DNA segments corresponding to the kinase insert domains of type III receptor tyrosine kinase genes in an endothelial cell library designated HL10246 (Clontech Laboratories, Inc., Palo Alto, CA). Degenerate oligonucleotide primers are designed which are complementary to conserved tyrosine kinase domains flanking the kinase insert domains of known type III receptor tyrosine kinases. These primers are used in the PCR procedure. DNA probes, designed from the DNA sequence of the PCR product, are then used to identify cDNA clones of the receptor gene from the original cDNA library.

In particular, the present invention relates to specific oligonucleotides which, when used as primers for PCR, allow for the amplification of DNA segments corresponding to the kinase insert domains of type III RTK genes.

In a principal embodiment, the present invention is directed to three overlapping DNA segments (designated BTIII081.8, BTIII129.5 and BTIV169) which comprise the entire coding region of this novel gene, namely, 4,068 nucleotides extending to the 3' end.

These DNA segments are isolated from a human endothelial cell cDNA library and together comprise the gene coding for a novel type III receptor tyrosine kinase. The human gene containing these DNA segments is referred to hereinafter as KDR (which stands for Kinase insert Domain containing Receptor) or, alternatively, as kdp (which stands for Kinase insert Domain containing Protein). The us of the term KDR is intend d to include any DNA s gments which form the

human gene which encodes the novel type III RTK of this application.

The DNA segments embodied in this invention are isolated from human sources. The present invention comprises DNA segments, and methods for using these DNA segments, which allow for the identification of a closely related gene in mouse DNA. The methods developed in this invention can be readily used by those skilled in the art for the identification and isolation of closely-related homologues in other species. Therefore, the present invention also embodies all DNA segments from species other than human which encode proteins having substantially the same amino acid sequence as that encoded by the kdp gene.

The present invention further relates to methods developed for the detection of mRNA's produced as a result of transcription of the sense strands of the DNA segments of this invention. Messenger RNA prepared from bovine endothelial cells are used in developing these methods. The ability to detect mRNA for a novel RTK may ultimately have medical benefit, especially in light of recent observations that the mRNA for certain RTKs are overexpressed in some cancers (5).

The methods developed in the present invention for detecting mRNA expressed by the kdp gene can be readily used by those of ordinary skill in the art for the detection of mRNA species related to the kdp gene in any cell type and from any species. For this reason, the present invention embodies all mRNA segments which are the result of transcription of the kdp gene.

The pr sent invention relates to m thods for expr ssion of the rec ptor prot in, for example, in CMT-3 cells of monk y kidney origin. Th receptor

protein, portions thereof, and mutated forms of the receptor protein may be expressed in many other cells by those skilled in the art using methods similar to those described in this application. For this reason, the present invention embodies all proteins encoded by the human KDR gene and proteins encoded by related genes found in other species.

The present invention further relates to methods for studying the interaction of VEGF to the expressed KDR protein. Recent work in the literature (17) indicates that VEGF is one member of a family of related proteins, and the interaction of growth factors similar to VEGF with the KDR protein can readily be studied by those skilled in the art using methods similar to those described in this application. These methods can readily be modified to study the interaction of candidate pharmaceuticals with the KDR protein towards the goal of developing an antagonist or agonist of VEGF action. For this reason, the present invention embodies methods for studying the interaction of VEGF and VEGF-related growth factors with the KDR protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a schematic representation of three receptor tyrosine kinase subclasses (6). KI is kinase insert domain; PTK is kinase domain; cys is cysteine rich region.

Figure 2 depicts the two sets of primers used for PCR (SEQ ID No: 1 and 2). The nucleotide sequences in appropriate regions of the four known type III r c ptor tyrosin kinas cDNAs are aligned and d gen rate oligonucl otide primers ar designed based upon th consensus s quences.

Figure 3 depicts the amplification of the kinase insert domains using PCR. DNA segments encoding the kinase insert domains of type III receptor tyrosine kinases are amplified by PCR. A sample (5 µl) is run on a 1.0% agarose gel which is stained with ethidium bromide. DNA size standards (123 bp ladder; Bethesda Research Laboratories, Bethesda, MD) are run as well.

Figure 4 depicts the DNA sequence of the two PCR products (Panel A: 363 bp segment derived from the 420 bp product (SEQ ID NO: 3); Panel B: 251 bp product (SEQ ID NO: 4)). The two products are purified by agarose gel electrophoresis, digested with Sall and EcoRI, and cloned into the plasmid vector pBlueScribe(+) (Strategene; San Diego, CA). The 420 bp PCR product is digested to 363 bp during this procedure. The DNA sequences for the primers used in the amplification are underlined.

Figure 5A depicts a computer assisted comparison of the DNA sequence for the 363 bp DNA segment derived from the 420 bp PCR product with the sequence of a DNA segment of the PDGF receptor (SEQ ID NO: 5) (18). A region of strong homology between the 363 bp segment derived from the 420 bp PCR product and the PDGF receptor is contained in a box. Figure 5B depicts a computer assisted comparison of the DNA sequence for the 251 bp PCR product with the sequence of a DNA segment of the FGF receptor (SEQ ID NO: 6) (7).

Figure 6 depicts the strategy used for sequencing the insert portions of clones BTIII081.8 and BTIII129.5 and BTIV169. The sequencing reaction uses either synthetic oligonucleotides (represented by boxes at the start of an arrow), or the M13 universal primer (no box) to initiate the reaction. In som cases, portions of th s DNA segments ar isolated using the

restriction enzymes indicated in the figure, and subcloned back into the plasmid vector pUC118, so that the M13 universal primer can be used. The position of the stop codon in BTIII129.5 is indicated. The coding portions of these DNA segments are shown at the bottom of the figure. The relative positions of the 1) membrane spanning portion, 2) kinase domains, and 3) kinase insert domain are indicated. The position of these structural features within the KDR derived DNA segments is compared in relation to their position in the PDGF-receptor ("PDGF-R").

Figure 7 depicts the DNA and predicted amino acid sequence of KDR, plus the stop codon (nucleotides 1-4071 of SEQ ID NO. 7). The sequence of the DNA segment amplified by PCR is underlined (nucleotides 2749-3105 of SEQ ID NO. 7). Cysteine residues in the putative extracellular domain are circled. Potential N-linked glycosylation sites are indicated by an asterisk. The putative membrane spanning region is enclosed in a box (nucleotides 2293-2367 of SEQ ID NO. 7).

Figure 8 depicts a hydropathy plot of the predicted amino acid sequence for the KDR protein.

Figure 9 depicts a comparison of the predicted amino acid sequence in the putative intracellular portion of the KDR protein to the ckit proto-oncogene (SEQ ID No: 8) (3), the CSF-1 receptor (SEQ ID NO: 9) (4), and the PDGF receptor (SEQ ID NO: 10) (18). Exact matches are indicated by an asterisk. Gaps are introduced to achieve maximum alignment. The putative ATP recognition site is indicated by three asterisks.

Figure 10 depicts th identification of kdp rec ptor mRNA by North rn blot analysis. Five micrograms of bovine aortic endothelial cell polyA+ RNA

are used. A nick-translated [^{32}P] CTP-labelled EcoRI/BamHI DNA segment (nucleotides 1510-2417 of SEQ ID NO. 7) is used as a probe. Autoradiography is for 36 hours.

Figure 11 depicts the kdp gene in human and mouse DNA by Southern blot analysis. A nick translated [^{32}P]CTP-labelled EcoRI/BamHI DNA segment (nucleotides 1510-2417 of SEQ ID NO. 7) is used as the probe. The probe is hybridized to Southern blots containing EcoRI digested DNA from human (lane 1), mouse (lane 2), and human-mouse hybrid cells (19) (lanes 3 and 4). The DNA used in lane 3 lacks the kdp locus, while DNA used in lane 4 contains the kdp locus.

Figure 12 depicts a Western blot analysis of CMT-3 cells which express the KDR protein. Cells are transfected with either the pcDNA1tkpASP vector alone (lane 1) or with that vector modified to contain the KDR gene (lane 2). 2 x 10^5 cells and 1 microgram of DNA are used for each transfection. Forty-eight hours later, Western blot analysis is performed on the samples using the anti-KDR.PS23 polyclonal antibody at a dilution of 1:1000. Detection of reacting proteins is performed using an ECL system (Amersham, Chicago, IL).

Figure 13 depicts the results of [^{125}I] VEGF binding to CMT-3 cells which express the KDR protein. Cells are transfected with either the vector alone (bars 1 and 2) or with the vector containing the KDR gene (bars 3 and 4). Forty-eight hours later, the samples are washed with phosphate buffered saline (PBS), and incubated with serum-free media containing 50 pM [^{125}I] VEGF (specific activity equal to 4,000 cpm per fmol), for 90 minutes. Nonradioactive VEGF, 5 nM, is added to some samples (bars 2 and 4) to defin

specific binding sites. The samples are washed with ice cold PBS, and the cells are transferred to gamma-counting tubes using 0.1% lubrol.

Figure 14 depicts the results of affinity cross-linking of [^{125}I] VEGF to CMT-3 cells which express the KDR protein. CMT-3 cells are transfected with either the vector alone (lane 1) or with the vector containing the KDR gene (lane 2). Forty-eight hours later, the cells are washed in PBS, and serum free media containing 200 pM [^{125}I] VEGF is added. After 90 minutes at room termperature, an affinity cross-linker disuccinimidyl suberate, 0.5 mM, is added for 15 minutes. The samples are then prepared for SDS-PAGE autoradiography.

DETAILED DESCRIPTION OF THE INVENTION

The strategy used to discover the DNA segments for the novel type III RTK gene begins with the design of two degenerate oligonucleotide primers based upon their homology to specific regions of the kinase domains of known RTK genes (Fig. 2) (3,4,7,18). In one embodiment, the polymerase chain reaction is then used to amplify DNA segments from a human endothelial cell cDNA library (designated HL 10246). The cDNA products from this step are each cloned into a plasmid vector designated pBlueScribe+ (Strategene, San Diego, CA) and sequenced. Oligonucleotide probes are designed from potentially interesting sequences in order to screen the cDNA library for more full length clones of the novel cDNA.

The strategy just described provides several novel elements: 1) the DNA sequences of the oligonuc-l otid primers us d during PCR; 2) the DNA s quence of the products generated by th polymerase chain

reaction; and 3) the DNA sequence of the final cloned DNA segments. Each of these elements of the invention described in this application will now be discussed in detail.

Figure 2 shows the rationale for choosing the oligonucleotide primers used in the PCR. The primers are designed to allow for the PCR amplification of the kinase insert domain of type III RTK genes. In order to design the primers, the DNA sequences of known type III RTK genes are aligned in specific regions of their catalytic domains, and a consensus sequence is chosen. The regions of the catalytic domains chosen in designing the primers flank the kinase insert domains of the receptor genes.

Primer 1 (SEQ ID No: 1) is designed from a region of the kinase domain 5' to the kinase insert domain, and consists of a mixture of four different 21mers. Primer 2 (SEQ ID NO: 2) is designed from a region of the kinase domain 3' to the kinase insert domain, and consists of a mixture of sixteen different 29mers with one inosine, indicated in SEQ ID NO: 2 by "N".

_Sal_I and _Eco_RI restriction sites are included at the 5' end of primers 1 and 2, respectively, to facilitate the subcloning of the amplified PCR products into plasmid vectors. Those skilled in the art may use other restriction sites; other minor modifications in the protocol above permits the design of primers without the inclusion of restriction sites.

The selection of these specific primers constitutes a novel approach towards identifying novel type III RTK genes. It had previously been shown (10) that primers designed from DNA sequences common to all tyr sine kinas s allows f r the id ntification f novel pr t ins. Th pres nt invention is th first to

contemplate the use of PCR to specifically target type III RTK.

The protocol used for PCR is as follows: Human endothelial cell cDNA (designated HL10246) is denatured by boiling and submitted to 30 cycles of PCR using 1 nmol of both primers in a final volume of 100 µl. The timing is 1.5 minutes at 92°C, 2 minutes at 50°C, and 2 minutes at 74°C. DNA from 5µl of sample is separated on a 1% agarose gel and stained with ethidium bromide.

Figure 3 shows the results of the PCR amplification. Two DNA products, with sizes 251 bp (SEQ ID NO: 4) and 420 bp, are visible when a sample of the reaction is electrophoresed on a 1.0% agarose gel and stained with ethidium bromide. The sizes of the two products are within the range expected for type III RTK genes (products derived from the FGF and PDGF receptor genes, which have the smallest and largest known kinase insert domains, would be 230 and 510 bp, respectively (20, 21).

The DNA from four continguous lanes with sizes ranging from 200 to 600 bp is electrophoresed onto DEAE filter paper, eluted from the paper with salt, and ethanol precipitated. The samples are incubated with 5 units of <u>Eco</u>RI and <u>Sal</u>I. The restriction enzymes digest the 420 bp DNA segment to a 363 bp DNA segment (SEQ ID NO: 3), due to the presence of an <u>Eco</u>RI site within the 420 bp DNA segment (nucleotide 2749, SEQ ID NO. 7). The restriction enzyme digested PCR products are then subcloned into the plasmid vector pBlueScribe(+). The recombinant clones are analyzed by sequencing using the dide xy-method (22) using a United States Biochemical (Cleveland, Ohio) S qu nase Version 2.0 sequencing kit. Figure 4 shows th DNA sequences for th 251 bp PCR

product and the 363 bp DNA segment derived from the 420 bp PCR product.

Computer assisted comparison of the DNA sequence for the 363 bp segment of the 420 bp PCR product to databases of known DNA sequences reveals that the sequence is novel, because it shares strong sequence identity with the flanking catalytic domain of known type III RTK genes, but not their kinase insert domains. Figure 5A compares the DNA sequence for the 363 DNA segment with that for the PDGF receptor gene (SEQ ID No: 5). Similar results are obtained using other type III RTK genes.

DNA sequencing of the 251 bp PCR product reveals a novel sequence containing both primers used for the amplification, but the sequence shows little homology to known tyrosine kinases. This is depicted in Figure 5B, which compares the DNA sequence for the 251 bp DNA segment with that for the FGF receptor (SEQ ID NO: 6). For this reason, further analysis of Product 1 is not pursued.

The protocols used during the PCR do not allow for amplification of the kinase insert domains of known receptor tyrosine kinases in the endothelial cell library used because of the low copy number of the message present in the library. There have been many studies on the effect of FGF on endothelial cell function (23,24) although there is evidence that the expression of the FGF receptor is developmentally regulated (7) and it is likely that the library used contains little or no cDNA for the FGF receptor.

An oligonucleotide probe, designed from the DNA sequence of the 363 bp segment, is synthesized (using an ABI 380 DNA Synthesizer) in order to screen the human endothelial cell cDNA library (HL10246) for the isolation of more full l ngth clones containing the

363 bp DNA segment. The probe sequence is chosen from the region of the 363 bp DNA segment which shares little sequence homology with known RTK.

The screening of the endothelial cell cDNA library is conducted as follows: Lambda gt11 phage, 10^6, are adsorbed to E. coli LE392 for 15 minutes at 37°C prior to plating onto agar plates at a density of 5×10^5 phage per plate. After allowing the phage plaques to develop at 37°C, plaque lifts are made using nitrocellulose filters, denatured in 0.4 N NaCl for 1 minute, and neutralized in 0.5 M Tris.HCl, pH 7.3, plus 1.5 M NaCl. The filters are washed with 2 x standard saline citrate (SSC) and then baked for 1.5 hour in a vacuum oven at 80°C. The filters are probed with an [^{32}P] ATP end labeled synthetic oligonucleotide, 5' - TTTCCCTTGACGGAATCGTGCCCCTTTGGT-3', which is the reverse complement of a DNA sequence contained in the PCR amplified product (Fig. 3). Hybridization is performed at 50°C in 5 x SSPE (167 mM NaCl, 10 mM sodium phosphate, pH 7.4, 1 mM EDTA), 2.5 x Denhardts, 0.5% sodium dodecyl sulfate (SDS), 100 µg/ml salmon sperm DNA. The filters are washed twice, 20 minutes per wash, with 2 x SSC plus 0.1% SDS at room temperature, followed by washing twice at 50°C with 0.1 X SSC plus 0.1% SDS; 20 minutes per wash. Positive clones are identified, picked and plaque purified.

Forty-five positive clones are obtained. Three of these positive clones are plaque purified and their phage DNA isolated. Digestion of the DNA with EcoRI and electrophoresis in agarose indicates that one clone, designated BTIII081.8, contains the largest insert, and subsequent analysis indicates that the DNA insert of this clone overlaps that of the inserts contained in other two purifi d clones (designat d BTIII079.11 and BTIII079.47A).

Digestion of the purified phage DNA of the clone designated BTIII081.8 with EcoRI results in DNA segments of 250 bp, 600 bp, and 1000 bp. Each of these three products is subcloned into the plasmid vector pUC118 and sequenced (Figure 6 shows the strategy used for sequencing). The orientation of the three fragments is determined by subcloning from the insert a BglII/BglII fragment into pUC118 and sequencing across the EcoRI junctions using a synthetic oligonucleotide to prime the sequencing reaction.

A restriction map is determined for each fragment (Figure 6). Various restriction site pieces are removed from the plasmids and recloned into pUC118 so that sequencing the resulting plasmids with the universal primer allows for sequencing most of the entire original fragments in both directions. Three oligonucleotide primers are required to sequence the entire cDNA in both directions. For the purposes of this application, this insert contains nucleotides numbered 1510-3406 (SEQ ID NO. 7).

A [^{32}P]CTP-labelled, nick-translated EcoRI-BamHI DNA segment derived from clone BTIII081.8 (nucleotides 1510-2417 of SEQ ID NO. 7) is used as a probe to rescreen the original endothelial cell cDNA library for more 5' full length DNA segments of the gene from which the insert portion of BTIII081.8 is derived. The protocols used to isolate the overlapping clones are identical to that used to isolate BTIII081.8.

A synthetic oligonucleotide probe is designed with 29 nucleotides corresponding to part of the DNA sequence of the insert portion of the clone BTIII081.8 (nucleotides 3297-3325 of SEQ ID NO. 7) in order to r screen the original ndoth lial cell cDNA library for more full 3' length DNA segm nts of the gene from which

the insert portion of BTIII081.8 is derived. The protocols used to isolate the overlapping clones are identical to that used to isolate BTIII081.8. Several positive clones for each of the 5' and 3' ends are identified and plaque purified.

One of the clones is designated BTIII200.2. The DNA from BTIII200.2 contains a 3.4 kb insert as determined by EcoRI digestion of the isolated phage DNA. EcoRI digestion of BTIII200.2 results in three DNA fragments. One of thse fragments (2.5 kb) is cloned into pUC119 and is designated BTIV006. The clone BTIV006 contains nucleotides numbered 7-2482. As described below, BTIV006 plus nucleotides 1-6 is designated BTIV169. DNA sequencing of the 2.5 kb DNA insert (BTIV169) indicates that it overlaps over one thousand nucleotides of the DNA sequence of the insert portion of the clone BTIII081.8 (Figure 6) at the 5' end.

A second clone isolated from the cDNA library is designated BTIII129.5. The DNA from BTIII129.5 contains a 2.2 kb insert as determined by EcoRI digestion of the isolated phage DNA. DNA sequencing of the 2.2 kb DNA insert indicates that it overlaps over five hundred nucleotides of the DNA sequence of the insert portion of the clone BTIII081.8 (Figure 6). The clone BTIII129.5 contains nucleotides numbered 2848-4236 (SEQ ID NO. 7). The DNA sequence for BTIII129.5 contains the stop codon TAA, defining the position of the 3' end of an open reading frame for the novel gene. Except for the first six nucleotides of the gene which are discussed below, these three clones define a gene encoding a growth factor receptor. These thr e clones define a 4,062 nucl otide s quenc of the open reading frame of the gene ext nding to the 3' end, follow d by a 168 nucleotid non-coding region (SEQ ID

NO. 7). A sample of a lambda gt11 phage harboring the clone BTIII081.8 has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A., and has been assigned ATCC accession number 40,931. A sample of a lambda gt11 phage harboring the clone BTIII129.5 has been deposited with the American Type Culture Collection and has been assigned ATCC accession number 40,975. For reasons discussed below, a sample of the clone BTIV006 was not deposited.

The aforementioned DNA segments (BTIII081.8, BTIII129.5, and BTIII200.2 (or BTIV006) encode 4062 nucleotides of the coding portion of a novel gene. The cDNA clones are incomplete in that a transcription initiation coding for methionine is missing. After the isolation of these clones, Matthews et al. (25) reported the cloning of a gene homologue of KDR in mouse, which was referred to as Flk-1. Analysis of the nucleic acid and amino acid sequence of Flk-1 indicated that the addition of six nucleotides to the 5' end of the isolated KDR clones would provide for a complete coding region.

To achieve this, an EcoRI-BamHI restriction fragment of BTIV200.2 is cloned into the plasmid pBlueScript KS (Strategene, La Jolla, CA). The 5' end of the inserted DNA is blunt ended with Klenow polymerase and Mung Bean nuclease. Next, the synthetic oligonucleotide TCGACGCGCG ATG GAG (SEQ ID NO. 11) is cloned into this vector. The oligonucleotide contains the sequence ATG GAG in frame with the downstream DNA insert. These nucleotides (ATG GAG) encode the amino acids methionine and glutamic acid, the first two amino acids encoded by the KDR g n . The resulting plasmid v ctor is designated BTIV140. This plasmid is purifi d on a CsCl gradient.

The purified plasmid is designated BTIV169. The insert of BTIV169 contains nucleotides 1-2400 (SEQ ID NO. 7) of the KDR gene. A sample of the plasmid pBlueScript KS which contains the clone BTIV169 has been deposited with the American Type Culture Collection and has been assigned ATCC accession number 75200.

Thus, together the clones BTIII081.8, BTIII129.5 and BTIV169 comprise the entire open reading frame of 4,068 nucleotides for the novel KDR gene. As will be discussed below, the KDR gene expresses the novel KDR receptor which binds specifically to the growth factor VEGF.

DNA sequencing of BTIII081.8, BTIII129.5 and BTIV169 (SEQ ID NO. 7) shows that the newly isolated gene is similar to, but distinct from, previously identified type III RTK. The predicted amino acid sequence (SEQ ID NO. 7) contains several structural features which demonstrate that the novel gene is a type III RTK. These structural features are summarized as follows:

1) A hydropathy plot of the predicted amino acid sequence indicates a single membrane spanning region (see Figure 8). This is characteristic of a type III RTK (Figure 7).

2) The putative amino-terminal 762 amino acid portion of the receptor has structural features of extracellular receptor ligand binding domains (1), including regularly spaced cysteines and 18 potential N-linked glycosylation sites (Figure 7).

3) The predicted amino acid sequence of the carboxy-terminal 530 amino acid portion contains an ATP-binding sit at lysine 868, 22 amino acids downstream from th consensus ATP r cognition sequence Gly-X-Gly-X-X-Gly (26) (Figur 8).

4) Within the kinase domain there is a 55-60% identical match in amino acid sequence to three other type III receptor tyrosine kinases: ckit proto-oncogene (SEQ ID NO: 8), CSF-1 (SEQ ID NO: 9) and PDGF (SEQ ID NO: 10) (Figure 9).

5) The predicted kinase domain contains a kinase insert domain of approximately 71 amino acids. As indicated in Figure 9, this portion of the amino acid sequence shares little sequence homology with other type III RTK.

The endothelial cell library can be further screened to isolate the 5' untranslated region and genomic clones can be generated so as to isolate the promoter region for the KDR gene.

In addition to the DNA sequence described for the KDR gene (SEQ ID NO. 7), the present invention further comprises DNA sequences which, by virtue of the redundancy of the genetic code, are biologically equivalent to the sequences which encode for the receptor, that is, these other DNA sequences are characterized by nucleotide sequences which differ from those set forth herein, but which encode a receptor having the same amino acid sequences as those encoded by the DNA sequences set forth herein.

In particular, the invention contemplates those DNA sequences which are sufficiently duplicative of the sequence of SEQ ID NO. 7 so as to permit hybridization therewith under standard high stringency Southern hybridization conditions, such as those described in Sambrook et al. (27), as well as the biologically active proteins produced thereby.

This invention also comprises DNA sequences which encod amino acid s quenc s which differ from those of th novel r ceptor, but which are the biological equivalent to thos d scribed for th

receptor. Such amino acid sequences may be said to be biologically equivalent to those of the receptor if their sequences differ only by minor deletions from or conservative substitutions to the receptor sequence, such that the tertiary configurations of the sequences are essentially unchanged from those of the receptor.

For example, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, as well as changes based on similarities of residues in their hydropathic index, can also be expected to produce a biologically equivalent product. Nucleotide changes which result in alteration of the N-terminal or C-terminal portions of the protein molecule would also not be expected to alter the activity of the protein. It may also be desirable to eliminate one or more of the cysteines present in the sequence, as the presence of cysteines may result in the undesirable formation of multimers when the protein is produced recombinantly, thereby complicating the purification and crystallization processes. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products. Therefore, where the terms "KDR gene" or "KDR protein" are used in either the specification or the claims, each will be understood to encompass all such modifications and variations which r sult in th production of a biologically quivalent protein.

In addition to the full length gene and protein, the invention encompasses biologically active fragments of each. By "biologically active" is meant a protein fragment which qualitatively retains the receptor activity of the larger KDR protein, or, in the case of a nucleotide sequence, which encodes such a protein fragment. It also refers, for purposes of antibody production, to fragments which are capable of eliciting production of antibodies capable of binding to the receptor protein.

To determine the size of the mRNA transcribed from the kdp gene, Northern blot hybridization experiments are carried out using an EcoRI/BamHI DNA segment (nucleotides 1510-2417, SEQ ID NO. 7) as a hybridization probe. The DNA used for the probe does not contain any portion of the putative kinase domain, and shares little sequence homology to other tyrosine kinases. The Northern blot analysis (Figure 10) shows that a 7 kb band is visualized in cytoplasmic poly(A)+ RNA of ABAE bovine aortic endothelial cells. This transcript differs in size from previously reported transcripts for known type III RTK (7,18).

The isolated cDNA is significant for several reasons. The cDNA encodes a novel type III receptor tyrosine kinase. The homology between the sequence of this cDNA and that of other receptors, as well as structural properties implied by the predicted amino acid sequence confirm the relationship. Receptors for growth factors should have tremendous utility in drug development as they face the outside of the cell and thus are among the best targets for drugs. In addition, the cellular levels of some receptors, in particular the neu proto-oncogen , increase during some canc rs. This has b en taken advantage of in designing diagnostic tests for these canc rs.

Southern analysis demonstrates that the kdp gene is present in mouse as well as human DNA. Mouse and human (Hela cell) DNA, 15 μg of each, are digested with 10 units of EcoRI and electrophoresed on a 0.7% agarose gel. The DNA is transferred onto nitrocellulose. The filter is hybridized to a [^{32}P]CTP-labelled cDNA probe made by nick translating an EcoRI/BamHI fragment from the 5' end of the kdp cDNA (nucleotides 1510-2417, SEQ ID NO. 7). Hybridization is conducted at 30°C in 5 X SSPE, 50% formamide, 0.1% SDS, plus 150 μg/ml salmon sperm DNA. The DNA probe hybridizes to Southern blots containing EcoRI digested DNA. After 48 hours, the filter is washed at room temperature in 2 X SSC plus 0.1% SDS for 20 minutes, followed by two 20 minute washes at 40°C with 0.1 X SSC plus 0.1% SDS. Autoradiography is then performed for 48 hours. As shown in Figure 11, radioactively labelled DNA is present in both human and mouse samples. This indicates that the kdp gene is present in both species.

An experiment is conducted to ascertain the genetic locus of kdp on human chromosomes. Thirty-eight cell hybrids from 18 unrelated human cell lines and four mouse cell lines are examined (19). A DNA probe hybridizes to Southern blots which contain EcoRI digested DNA from the human-mouse hybrids (using the procedure and DNA probe for human and mouse tissue described in relation to Figure 11). Table I sets forth the results of the segregation of kdp with human chromosomes in EcoRI digested human-mouse somatic cell hybrid DNA:

Table I

Chromosome	Concordant # of Hybrids (+/+)	(-/-)	Discordant # of Hybrids (+/-)	(-/+)	% Discordancy
1	4	19	8	4	34
2	8	18	5	6	30
3	11	12	3	9	34
4	14	24	0	0	0
5	7	14	7	10	45
6	7	19	7	5	32
7	11	14	3	8	31
8	8	11	6	13	50
9	3	20	10	4	38
10	12	9	2	14	43
11	9	13	4	11	41
12	9	10	5	14	50
13	7	18	7	6	34
14	11	8	3	16	50
15	9	15	5	8	35
16	7	19	7	5	32
17	12	7	2	16	49
18	11	14	3	10	34
19	7	18	7	6	34
20	9	10	5	14	50
21	11	9	3	15	47
22	3	16	10	7	47
X	8	10	3	8	38

The scoring is determined by the presence(+) or absence (-) of human bands in the hybrids on Southern blots prepared in a similar to those shown in Figure 11. Th sc ring is compar d to the presence or abs nce of human chr mos m s in each hybrid. A 0%

discordancy indicates a matched segregation of the DNA probe with a chromosome. Three fragments, approximately 6.5 kb, 3.1 kb, and 0.7 kb in size are detected in digests of human DNA (Figure 11), and in all hybrids which had retained human chromosome 4 (Table I). All other chromosomes are excluded in at least 11 discordant hybrids (Table I). The results of Figure 11 and Table I demonstrate that the genetic locus of kdp is on human chromosome 4.

It is noteworthy that both the ckit (3) and the type A PDGF (28) receptor genes map to human chromosome 4. The finding that the genetic locus of kdp is on human chromosome 4 provides further evidence that the novel receptor of this invention is a type III receptor tyrosine kinase.

The next step after identifying the entire coding portion of the kdp gene is to express the receptor protein encoded by that gene. The receptor protein is then utilized so as to identify the growth factor which binds specifically to the receptor.

The receptor protein is expressed using established recombinant DNA methods. Suitable host organisms include bacteria, viruses, yeast, insect or mammalian cell lines, as well as other conventional organisms. For example, CMT-3 monkey kidney cells are tranfected with a vector containing the complete coding region of the KDR gene.

The complete coding portion of the KDR gene is assembled by sequentially cloning into pUC119 three DNA fragments derived from BTIII081.8, BTIII129.5, and BTIV169. First, a SmaI-EcoRI fragment of clone BTIII129.5 (nucleotides 3152-4236, SEQ ID NO. 7) is blunt end d with Klenow polym rase and introduc d into a SmaI sit in pUC119. Next, a BamHI-SmaI fragment of clon BTIII081.8 (nucl otid s 2418-3151, SEQ ID NO. 7)

is introduced at a BamHI-SmaI site. Finally, a SalI-BamHI fragment of clone BTIV169 (nucleotides 1-2417, SEQ ID NO. 7) is introduced at a SalI-BamHI site. Part of the cloning site of pUC119 is contained in the SalI-BamHI fragment, 5' to the KDR gene. In order to clone the complete coding portion into an expression vector, the assembled DNA (in pUC119) is digested with SalI and Asp118 and recloned into the eukaryotic expression vector pcDNA1tkpASP.

This vector is a modification of the vector pcDNA1 (Invitrogen; San Diego, CA). Specifically, the ampicillin resistance gene is cloned from pBR322 into pcDNA1. A small SV40 T splice and the SV40 polyadenylation signal are then removed and are replaced with a Herpes Simplex Virus-1 polyadenylation signal. Finally, a cytomegalovirus intermediate early splice is inserted 5' to the cloning site to yield pcDNA1tkpASP.

Transfection of CMT-3 cells is done using DEAE-dextran. Forty-eight hours after transfection, expression of the novel receptor is monitored using Western blot analysis as follows.

An antibody is used to assay the expressed receptor protein. The predicted amino acid sequence of the receptor is used to generate peptide-derived antibodies to the receptor by conventional techniques. The presence of the novel receptor protein is confirmed by Western blot hybridization.

Specifically, a synthetic peptide with 13 residues is synthesized based on the 12 residues corresponding to amino acids 986-997 of the putative amino acid sequence of the KDR protein (SEQ ID NO. 7), with a cysteine residue linked to the lysine (amino acid 997). The cysteine facilitates coupling of the p ptide to a macromol cul which functions as a carrier for the peptid . For xample, th peptide is coupled

to keyhole limpet haemocyanin (KLH) using m-maleimidobenzoyl-N-hydroxysuccinimide ester. Other conventional carriers may be used such as human and bovine serum albumins, myoglobins, β-galactosidase, penicillinase and bacterial toxoids, as well as synthetic molecules such as multi-poly-DL-alanyl-poly-L-lysine and poly-L-lysine.

Rabbits are immunized with the peptide-KLH conjugate to raise polyclonal antibodies. After different periods of time, serum is collected from the rabbits. The IgG fraction of the serum is then purified using a protein A Sepharose column (Pharmacia LKB, Uppsala, Sweden) to obtain the antibody which is designated anti-<u>KDR</u>.PS23.

A sample of the expressed <u>KDR</u> protein is subjected to SDS-PAGE using a 7% acrylamide gel under standard conditions. The protein band is then transferred on to nitrocellulose paper for Western blot analysis and the anti-<u>KDR</u>.PS23 antibody is added at a dilution of 1:1,000 to allow the antibody to react with the protein present. A second antibody, goat anti-rabbit antibody to rabbit IgG, which binds to anti-<u>KDR</u>.PS23, is then added. The detection of proteins which react with the antibodies is performed by autoradiography of bands using an ECL system (Amersham, Chicago, IL). The results are depicted in Figure 12.

Figure 12 shows that a 190 kD protein is present in the cells transfected with the vector containing the <u>KDR</u> gene, but is absent in cells transfected with vector alone. The size of this protein is consistent with it being encoded by the <u>KDR</u> g n , in that th pr dict d amino acid s quenc for the unglycosylat d <u>KDR</u> prot in is 156 kD, and that sequence c ntains 18 putativ extracellular glycosylation sit s

which would account for the balance of the size seen in the 190 kD band.

The expressed receptor is then used to identify the growth factor which interacts with the receptor. In order to test the hypothesis that the KDR protein is a receptor for VEGF, radioligand binding studies are performed. VEGF (provided by D. Gospodarowicz) is radiolabelled with ^{125}I. Cells are transfected with either the vector pcDNAltkpASP alone (bars 1 and 2 of Figure 13) or with the vector containing the KDR gene (bars 3 and 4). Forty-eight hours later, the transfected cell samples are washed with PBS and then incubated for 90 minutes with serum-free media containing 50 pM [^{125}I]VEGF (specific activity equal to 4,000 cpm per fmol). Excess nonradioactive VEGF, 5 nM, is added to some samples (bars 2 and 4) to define specific binding sites. The samples are washed with ice cold PBS, and the cells are transferred to gamma-counting tubes using a detergent, 0.1% lubrol.

The results of the radioligand binding studies are depicted in Figure 13. Figure 13 shows that CMT-3 cells transfected with vector containing the KDR gene contain specific binding sites for [^{125}I]VEGF (compare bars 3 and 4), while cells transfected with vector alone do not (compare bars 1 and 2).

Further evidence that the KDR gene encodes a receptor for VEGF is demonstrated by affinity cross-linking studies (Figure 14). Figure 14 depicts the results of affinity cross-linking of [^{125}I]VEGF to CMT-3 cells which express the KDR protein. CMT-3 cells are transfected with either the pcDNAltkpASP vector alone (lan 1 of Figure 14) or with the vector containing the KDR gene (lane 2). Forty-eight hours lat r, the cells are washed in PBS, and serum fr e

media containing 200 pM [^{125}I]VEGF is added. After 90 minutes at room temperature, an affinity cross-linker disuccinimidyl suberate (Pierce Biochemicals, Rockford, IL), 0.5mM, is added for 15 minutes. The samples are then subjected to SDS-PAGE autoradiography.

Three protein bands are seen in SDS-PAGE autoradiograms from samples of CMT-3 cells transfected with the KDR gene and cross-linked to [^{125}I]VEGF (lane 1). The size of band 1 (235 kD) is consistent with it being the 190 kD protein seen by Western blot analysis (Figure 12), because a 45 kD [^{125}I] VEGF dimer plus 190 kD would migrate in a manner identical to band 1. The origin of band 2 is not clear, but may represent an altered glycosylation form of band 1. Band 3 (22.5 kD) is most likely VEGF itself, and can be seen faintly in cells transfected with vector alone (lane 2).

The novel KDR gene of this invention is significant for several reasons. Studies of the cellular mechanisms by which receptors function in signal transduction have led in the past to a better understanding of how cells grow in both normal and diseased states. Receptor tyrosine kinases, in particular, have received a great deal of attention because of the observation that a number of RTK are the cellular counterparts for viral oncogenes, implying a direct correlation between changes in the expression of RTK and cancer. In view of this, it is likely that pharmaceuticals targeted at RTK will inhibit the changes in cell growth associated with cancer. In additon, it is likely that monitoring the levels of expression of RTK will prove valuable in diagnosing the onset of cancer.

The d scrib d cDNA is isolated from a human endothelial cell library. Endoth lial c lls participate in angiogenesis, the formation of new blood

capillaries. Previous work directed towards identifying the growth factors which regulate angiogenesis have primarily focused upon FGF (13), although recent evidence has indicated that other growth factors may be involved as well (12,15,29). This evidence consists of the observations that: 1) FGF does not contain a signal sequence (24) and thus may not be secreted from cells in a manner consistent with the tight regulation of angiogenesis, and 2) endothelial cells synthesize FGF and yet are normally resting (15). Our discovery, then, of a novel growth factor receptor may ultimately clarify these inconsistencies and lead to a better understanding of endothelial cell function.

The teachings of this invention can be readily used by those skilled the art for the purpose of testing pharmaceuticals targeted at the KDR protein. Two examples of approaches which can be used for this purpose are now given.

First, the methods described in this invention for studying the interaction of VEGF with KDR protein can be used to test for pharmaceuticals which will antagonize that interaction. For these studies, cells expressing the KDR protein are incubated with [^{125}I]VEGF, together with a candidate pharmaceutical. Inhibition of radioligand binding is tested for; significant inhibition indicates the candidate is an antagonist. Permanent expression of the KDR protein in a cell type such as NIH3T3 cells would make these studies less laborious. This can be easily achieved by those skilled in the art using the described methods.

Second, using the teachings of this inv ntion, those skill-d in the art can study structural properties of the KDR protein inv lved in r ceptor function. This structural information can

then be used to more rationally design pharmaceuticals which inhibit that function. Mutagenesis of the KDR gene by well established protocols is one approach, crystallization of the receptor binding site is another.

Bibliography

1. Yarden Y., and A. Ullrich, Ann. Rev. Biochem., 57, 433-478 (1988).

2. Bargmann, C., et al., Nature, 319, 226-230 (1986).

3. Yarden, Y., et al., EMBO J., 6, 3341-3351 (1987).

4. Coussens, L., et al., Nature, 320, 277-280 (1986).

5. Slamon, D., et al., Science, 244, 707-712 (1989).

6. Ullrich, A. and Schlessinger, J., Cell, 61, 203-212 (1990).

7. Ruta, M., et al., Oncogene, 3, 9-15 (1988).

8. Strathmann, M., et al., Proc. Natl. Acad. Sci., 86, 8698-8702 (1989).

9. Streuli, M., et al., Proc. Natl. Acad. Sci., 86, 8698-8702 (1989).

10. Wilkes, A.F., Proc. Natl. Acad. Sci., 86, 1603-1607 (1989).

11. Folkman, J., and Klagsbrun, M., Science, 235, 442-445 (1987).

12. Ishikawa, F., et al., Nature, 338, 557-562 (1989).

13. Baird, A., and Bohlen, P., in Peptide Growth Factors and Their Receptors, pages 369-418 (Spron, M.B., and Roberts, A.B., eds. 1990).

14. Senger, D.R., et al., Science, 219, 983-985 (1983).

15. Gospodarowicz, D., et al., Proc. Natl. Acad. Sci., 86, 7311-7315 (1989).

16. Leung, D.W., t al., Science, 246, 1306-1309 (1989).

17. Maglione, D., et al., Proc. Natl. Acad. Sci., 88, 9267-9271 (1991).

18. Gronwald, R., et al., Proc. Natl. Acad. Sci., 85, 3435-3439 (1988).

19. Shows, T., et al., Somat. Del. Mol. Gen., 10, 315-318 (1984).

20. Rainer, G., et al., Proc. Natl. Acad. Sci., 85, 3435-3439 (1988).

21. Lee, P. L., et al., Science, 245, 57-60 (1989).

22. Sanger, F., et al., Proc. Natl. Acad. Sci., 74, 5463-5467 (1977).

23. Folkman, J., Cancer Res., 46, 467-473 (1986).

24. Burgess, W. and Maciag, T., Ann. Rev. Biochem., 58, 575-606 (1989).

25. Matthews, W., et al., Proc. Natl. Acad. Sci., 88, 9026-9030 (1991).

26. Hannink, M. and Donoghue, D., Proc. Natl. Acad. Sci., 82, 7894-7898 (1985).

27. Sambrook, J., et al., Molecule Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

28. Matsui, T., et al., Science, 243, 800-804 (1989).

29. Conn, G., et al., Proc. Natl. Acad. Sci., 87, 2628-2632 (1990).

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Terman, Bruce I
Carrion, Miguel E

(ii) TITLE OF INVENTION: Identification of a Novel Human Growth Factor Receptor

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Alan M. Gordon
American Cyanamid Company

(B) STREET: 1937 West Main Street, P.O. Box 60

(C) CITY: Stamford

(D) STATE: Connecticut

(E) COUNTRY: USA

(F) ZIP: 06904

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC AT

(C) OPERATING SYSTEM: MS-DOS

(D) SOFTWARE: ASCII from IBM DW 4

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/657,236

(B) FILING DATE: February 22, 1991

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Gordon, Alan M.

(B) REGISTRATION NUMBER: 30,637

(C) REFERENCE/DOCKET NUMBER: 31,298-01

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 203 321 2719

(B) TELEFAX: 203 321 2971

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GTCGAC AAY CTG TTG GGR GCC TGC AAC 27

(2) INFORMATION FOR SEQ ID NO: 2 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GAATTC AG CAC KTT NCT RGC YGC CAG GTC TGY GTC 35

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 363 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAA TTC TGC AAA TTT GGA AAC CTG TCC ACT TAC CTG 36

AGG ACG AAG AGA AAT GAA TTT GTC CCC TAC AAG ACC 72

AAA GGG GCA CGA TTC CGT CAA GGG AAA GAC TAC GTT 108

GGA GCA ATC CCT GTG GAT CTG AAA CGG CGC TTG GAC 144

ACG CAT CAC CAG TAG CCA GAG CTC AGC CAG CTC TGG 180

ATT TGT GGA GGA GAA GTC CCT CAG TGA TGT AGA AGA 216

AGA GGA AGC TCC TGA AGA TCT GTA TAA GGA CTT CCT 252

GAC CTT GGA GCA TCT CAT CTG TTA CAG TTT CCA AGT 288

GGC TAA GGG CAT GGA GTT CTT GGC ATC GCG AAA GTG 324

TAT CCA CAG AGA CCT GGC AGC CAG GAA CGT GCT GAA 360

TTC 363

(2) INFORMATION FOR SEQ ID NO: 4 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTC GAC AAT CTG TTG GGG GCC TGC ACC ATC CCA ACA 36

TCC TGC TGC TCT ACA ACT ATT TTT ATG ACC GGA GGA 72

GGA TCT ACT TGA TTC TAG AGT ATG CCC CCC GCG GAG 108

CTC TAC AAG GAG CTG CAG AAG AGC TGC ACA TTT GAC 144

GAG CAG CGA ACA GCC ACG ATC ATG GAG GAG TTG GCA 180

GAT GCT CTA ATG TAC TGC CGT GGG AAG AAG GTG ATT 216

CAC AGA GAC CTG GCA GCC AGC AAC GTG CTG AAT TC 251

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 510 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: PDGF Receptor DNA

(B) LOCATION: Internal sequence

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Gronwald, R., et al.

(B) JOURNAL: Proc. Natl. Acad. Sci.

(C) VOLUME: 85

(D) PAGES: 3435-3439

(E) DATE: 1988

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AAC CTG TGG GGG CCT GCA CCA AAG GAG GAC CAT CTA 36

TAT CAT CTA TAT CAT CAC TGA GTA CTG CCG CTA CGG 72

AGA CCT GGT GGA CTA CCT GCA CCG CAA CAA ACA CAC 108

CTT CCT GCA GCA CCA CTC CGA CAA GCG CCG CCC GCC 144

CAG CGC GGA GCT CTA CAG CAA TGC TCT GCC CGT TGG 180

GCT CCC CCT GCC CAG CCA TGT GTC CTT GAC CGG GGG 216

AGA GCG ACG GTG GCT ACA TGG ACA TGA GCA AGG ACG 252

AGT CGG TGG ACT ATG TGC CCA TGC TGG ACA TGA AAG 288

GAG ACG TCA AAT AGC AGA CAT CGA GTC CTC CAA CTA 324

CAT GGC CCC TTA CGA TAA CTA CGT TCC CTC TGC CCC 360

TGA GAG GAC CTG CCG AGC AAC TTT GAT CAA CGA GTC 396

TCC AGT GCT AAG CTA CAT GGA CCT CGT GGG CTT CAG 432

CTA CCA GGT GGC CAA TGG CAT GGA GTT CTG GCC TCC 468

AAG AAC TGC GTC CAC AGA GAC CTG GCG GCT AGG AAC 504

GTC CTT 510

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 255 bas pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: FGF Receptor DNA

(B) LOCATION: Internal sequence

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Ruta, M., et al.

(B) JOURNAL: Oncogene

(C) VOLUME: 3

(D) PAGES: 9-15

(E) DATE: 1988

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
AAC CTG CTG GGG GCC TGC ACG CAG GAT GGT CCC TTG    36

TAT GTC ATC GTG GAG TAT GCC TCC AAG GGC AAC CTG    72

CGG GAG TAC CTG CAG ACC CGG AGG CCC CCA GGG CTG   108

GAA TAC TGC TAT AAC CCC AGC CAC AAC CCA GAG GAG   144
```

CAG CTC TCC TCC AAG GAC CTG GTG TCC TGC GCC TAC 180

CAG GAG GCC CGA GGC ATG GAG TAT CTG GCC TCC AAG 216

AAG TGC ATA CAC CGA GAC CTG GCA GCC AGG AAT GTC 252

CTG 255

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4236 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
    ATG GAG AGC AAG GTG CTG CTG GCC GTC GCC CTG     33
    Met Glu Ser Lys Val Leu Leu Ala Val Ala Leu
     1               5                  10

TGG CTC TGC GTG GAG ACC CGG GCC GCC TCT GTG GGT     69
Trp Leu Cys Val Glu Thr Arg Ala Ala Ser Val Gly
            15                  20

TTG CCT AGT GTT TCT CTT GAT CTG CCC AGG CTC AGC    105
Leu Pro Ser Val Ser Leu Asp Leu Pro Arg Leu Ser
     25                  30                  35
```

ATA CAA AAA GAC ATA CTT ACA ATT AAG GCT AAT ACA 141
Ile Gln Lys Asp Ile Leu Thr Ile Lys Ala Asn Thr
40 45

ACT CTT CAA ATT ACT TGC AGG GGA CAG AGG GAC TTG 177
Thr Leu Gln Ile Thr Cys Arg Gly Gln Arg Asp Leu
50 55

GAC TGG CTT TGG CCC AAT AAT CAG AGT GGC AGT GAG 213
Asp Trp Leu Trp Pro Asn Asn Gln Ser Gly Ser Glu
60 65 70

CAA AGG GTG GAG GTG ACT GAG TGC AGC GAT GGC CTC 249
Gln Arg Val Glu Val Thr Glu Cys Ser Asp Gly Leu
75 80

TTC TGT AAG ACA CTC ACA ATT CCA AAA GTG ATC GGA 285
Phe Cys Lys Thr Leu Thr Ile Pro Lys Val Ile Gly
85 90 95

AAT GAC ACT GGA GCC TAC AAG TGC TTC TAC CGG GAA 321
Asn Asp Thr Gly Ala Tyr Lys Cys Phe Tyr Arg Glu
100 105

ACT GAC TTG GCC TCG GTC ATT TAT GTC TAT GTT CAA 357
Thr Asp Leu Ala Ser Val Ile Tyr Val Tyr Val Gln
110 115

GAT TAC AGA TCT CCA TTT ATT GCT TCT GTT AGT GAC 393
Asp Tyr Arg Ser Pro Phe Ile Ala Ser Val Ser Asp
120 125 130

CAA CAT GGA GTC GTG TAC ATT ACT GAG AAC AAA AAC 429
Gln His Gly Val Val Tyr Il Thr Glu Asn Lys Asn
135 140

```
AAA ACT GTG GTG ATT CCA TGT CTC GGG TCC ATT TCA    465
Lys Thr Val Val Ile Pro Cys Leu Gly Ser Ile Ser
    145                 150                 155

AAT CTC AAC GTG TCA CTT TGT GCA AGA TAC CCA GAA    501
Asn Leu Asn Val Ser Leu Cys Ala Arg Tyr Pro Glu
                160                 165

AAG AGA TTT GTT CCT GAT GGT AAC AGA ATT TCC TGG    537
Lys Arg Phe Val Pro Asp Gly Asn Arg Ile Ser Trp
        170                 175

GAC AGC AAG AAG GGC TTT ACT ATT CCC AGC TAC ATG    573
Asp Ser Lys Lys Gly Phe Thr Ile Pro Ser Tyr Met
180                 185                 190

ATC AGC TAT GCT GGC ATG GTC TTC TGT GAA GCA AAA    609
Ile Ser Tyr Ala Gly Met Val Phe Cys Glu Ala Lys
            195                 200

ATT AAT GAT GAA AGT TAC CAG TCT ATT ATG TAC ATA    645
Ile Asn Asp Glu Ser Tyr Gln Ser Ile Met Tyr Ile
    205                 210                 215

GTT GTC GTT GTA GGG TAT AGG ATT TAT GAT GTG GTT    681
Val Val Val Val Gly Tyr Arg Ile Tyr Asp Val Val
                220                 225

CTG AGT CCG TCT CAT GGA ATT GAA CTA TCT GTT GGA    717
Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly
        230                 235

GAA AAG CTT GTC TTA AAT TGT ACA GCA AGA ACT GAA    753
Glu Lys Leu Val Leu Asn Cys Thr Ala Arg Thr Glu
240                 245                 250
```

```
CTA AAT GTG GGG ATT GAC TTC AAC TGG GAA TAC CCT   789
Leu Asn Val Gly Ile Asp Phe Asn Trp Glu Tyr Pro
            255                 260

TCT TCG AAG CAT CAG CAT AAG AAA CTT GTA AAC CGA   825
Ser Ser Lys His Gln His Lys Lys Leu Val Asn Arg
    265                 270                 275

GAC CTA AAA ACC CAG TCT GGG AGT GAG ATG AAG AAA   861
Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys
                280                 285

TTT TTG AGC ACC TTA ACT ATA GAT GGT GTA ACC CGG   897
Phe Leu Ser Thr Leu Thr Ile Asp Gly Val Thr Arg
        290                 295

AGT GAC CAA GGA TTG TAC ACC TGT GCA GCA TCC AGT   933
Ser Asp Gln Gly Leu Tyr Thr Cys Ala Ala Ser Ser
300                 305                 310

GGG CTG ATG ACC AAG AAG AAC AGC ACA TTT GTC AGG   969
Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val Arg
            315                 320

GTC CAT GAA AAA CCT TTT GTT GCT TTT GGA AGT GGC  1005
Val His Glu Lys Pro Phe Val Ala Phe Gly Ser Gly
    325                 330                 335

ATG GAA TCT CTG GTG GAA GCC ACG GTG GGG GAG CGT  1041
Met Glu Ser Leu Val Glu Ala Thr Val Gly Glu Arg
                340                 345

GTC AGA ATC CCT GCG AAG TAC CTT GGT TAC CCA CCC  1077
Val Arg Il  Pro Ala Lys Tyr Leu Gly Tyr Pro Pro
        350                 355
```

```
CCA GAA ATA AAA TGG TAT AAA AAT GGA ATA CCC CTT  1113
Pro Glu Ile Lys Trp Tyr Lys Asn Gly Ile Pro Leu
360                 365                 370

GAG TCC AAT CAC ACA ATT AAA GCG GGG CAT GTA CTG  1149
Glu Ser Asn His Thr Ile Lys Ala Gly His Val Leu
            375                 380

ACG ATT ATG GAA GTG AGT GAA AGA GAC ACA GGA AAT  1185
Thr Ile Met Glu Val Ser Glu Arg Asp Thr Gly Asn
    385                 390                 395

TAC ACT GTC ATC CTT ACC AAT CCC ATT TCA AAG GAG  1221
Tyr Thr Val Ile Leu Thr Asn Pro Ile Ser Lys Glu
                400                 405

AAG CAG AGC CAT GTG GTC TCT CTG GTT GTG TAT GTC  1257
Lys Gln Ser His Val Val Ser Leu Val Val Tyr Val
        410                 415

CCA CCC CAG ATT GGT GAG AAA TCT CTA ATC TCT CCT  1293
Pro Pro Gln Ile Gly Glu Lys Ser Leu Ile Ser Pro
420                 425                 430

GTG GAT TCC TAC CAG TAC GGC ACC ACT CAA ACG CTG  1329
Val Asp Ser Tyr Gln Tyr Gly Thr Thr Gln Thr Leu
            435                 440

ACA TGT ACG GTC TAT GCC ATT CCT CCC CCG CAT CAC  1365
Thr Cys Thr Val Tyr Ala Ile Pro Pro Pro His His
    445                 450                 455

ATC CAC TGG TAT TGG CAG TTG GAG GAA GAG TGC GCC  1401
Il  His Trp Tyr Trp Gln Leu Glu Glu Glu Cys Ala
                460                 465
```

AAC GAG CCC AGC CAA GCT GTC TCA GTG ACA AAC CCA 1437
Asn Glu Pro Ser Gln Ala Val Ser Val Thr Asn Pro
470 475

TAC CCT TGT GAA GAA TGG AGA AGT GTG GAG GAC TTC 1473
Tyr Pro Cys Glu Glu Trp Arg Ser Val Glu Asp Phe
480 485 490

CAG GGA GGA AAT AAA ATT GAA GTT AAT AAA AAT CAA 1509
Gln Gly Gly Asn Lys Ile Glu Val Asn Lys Asn Gln
495 500

TTT GCT CTA ATT GAA GGA AAA AAC AAA ACT GTA AGT 1545
Phe Ala Leu Ile Glu Gly Lys Asn Lys Thr Val Ser
505 510 515

ACC CTT GTT ATC CAA GCG GCA AAT GTG TCA GCT TTG 1581
Thr Leu Val Ile Gln Ala Ala Asn Val Ser Ala Leu
520 525

TAC AAA TGT GAA GCG GTC AAC AAA GTC GGG AGA GGA 1617
Tyr Lys Cys Glu Ala Val Asn Lys Val Gly Arg Gly
530 535

GAG AGG GTG ATC TCC TTC CAC GTG ACC AGG GGT CCT 1653
Glu Arg Val Ile Ser Phe His Val Thr Arg Gly Pro
540 545 550

GAA ATT ACT TTG CAA CCT GAC ATG CAG CCC ACT GAG 1689
Glu Ile Thr Leu Gln Pro Asp Met Gln Pro Thr Glu
555 560

CAG GAG AGC GTG TCT TTG TGG TGC ACT GCA GAC AGA 1725
Gln Glu Ser Val Ser Leu Trp Cys Thr Ala Asp Arg
565 570 575

```
TCT ACG TTT GAG AAC CTC ACA TGG TAC AAG CTT GGC  1761
Ser Thr Phe Glu Asn Leu Thr Trp Tyr Lys Leu Gly
                580                 585

CCA CAG CCT CTG CCA ATC CAT GTG GGA GAG TTG CCC  1797
Pro Gln Pro Leu Pro Ile His Val Gly Glu Leu Pro
        590                 595

ACA CCT GTT TGC AAG AAC TTG GAT ACT CTT TGG AAA  1833
Thr Pro Val Cys Lys Asn Leu Asp Thr Leu Trp Lys
600                 605                 610

TTG AAT GCC ACC ATG TTC TCT AAT AGC ACA AAT GAC  1869
Leu Asn Ala Thr Met Phe Ser Asn Ser Thr Asn Asp
            615                 620

ATT TTG ATC ATG GAG CTT AAG AAT GCA TCC TTG CAG  1905
Ile Leu Ile Met Glu Leu Lys Asn Ala Ser Leu Gln
    625                 630                 635

GAC CAA GGA GAC TAT GTC TGC CTT GCT CAA GAC AGG  1941
Asp Gln Gly Asp Tyr Val Cys Leu Ala Gln Asp Arg
                640                 645

AAG ACC AAG AAA AGA CAT TGC GTG GTC AGG CAG CTC  1977
Lys Thr Lys Lys Arg His Cys Val Val Arg Gln Leu
        650                 655

ACA GTC CTA GAG CGT GTG GCA CCC ACG ATC ACA GGA  2013
Thr Val Leu Glu Arg Val Ala Pro Thr Ile Thr Gly
660                 665                 670

AAC CTG GAG AAT CAG ACG ACA AGT ATT GGG GAA AGC  2049
Asn Leu Glu Asn Gln Thr Thr Ser Ile Gly Glu S r
            675                 680
```

ATC GAA GTC TCA TGC ACG GCA TCT GGG AAT CCC CCT 2085
Ile Glu Val Ser Cys Thr Ala Ser Gly Asn Pro Pro
685 690 695

CCA CAG ATC ATG TGG TTT AAA GAT AAT GAG ACC CTT 2121
Pro Gln Ile Met Trp Phe Lys Asp Asn Glu Thr Leu
700 705

GTA GAA GAC TCA GGC ATT GTA TTG AAG GAT GGG AAC 2157
Val Glu Asp Ser Gly Ile Val Leu Lys Asp Gly Asn
710 715

CGG AAC CTC ACT ATC CGC AGA GTG AGG AAG GAG GAC 2193
Arg Asn Leu Thr Ile Arg Arg Val Arg Lys Glu Asp
720 725 730

GAA GGC CTC TAC ACC TGC CAG GCA TGC AGT GTT CTT 2229
Glu Gly Leu Tyr Thr Cys Gln Ala Cys Ser Val Leu
735 740

GGC TGT GCA AAA GTG GAG GCA TTT TTC ATA ATA GAA 2265
Gly Cys Ala Lys Val Glu Ala Phe Phe Ile Ile Glu
745 750 755

GGT GCC CAG GAA AAG ACG AAC TTG GAA ATC ATT ATT 2301
Gly Ala Gln Glu Lys Thr Asn Leu Glu Ile Ile Ile
760 765

CTA GTA GGC ACG ACG GTG ATT GCC ATG TTC TTC TGG 2337
Leu Val Gly Thr Thr Val Ile Ala Met Phe Phe Trp
770 775

CTA CTT CTT GTC ATC ATC CTA GGG ACC GTT AAG CGG 2373
Leu Leu Leu Val Ile Il Leu Gly Thr Val Lys Arg
780 785 790

GCC AAT GGA GGG GAA CTG AAG ACA GGC TAC TTG TCC 2409
Ala Asn Gly Gly Glu Leu Lys Thr Gly Tyr Leu Ser
795 800

ATC GTC ATG GAT CCA GAT GAA CTC CCA TTG GAT GAA 2445
Ile Val Met Asp Pro Asp Glu Leu Pro Leu Asp Glu
805 810 815

CAT TGT GAA CGA CTG CCT TAT GAT GCC AGC AAA TGG 2481
His Cys Glu Arg Leu Pro Tyr Asp Ala Ser Lys Trp
820 825

GAA TTC CCC AGA GAC CGG CTG AAC CTA GGT AAG CCT 2517
Glu Phe Pro Arg Asp Arg Leu Asn Leu Gly Lys Pro
830 835

CTT GGC CGT GGT GCC TTT GGC CAA GAG ATT GAA GCA 2553
Leu Gly Arg Gly Ala Phe Gly Gln Glu Ile Glu Ala
840 845 850

GAT GCC TTT GGA ATT GAC AAG ACA GCA ACT TGC AGG 2589
Asp Ala Phe Gly Ile Asp Lys Thr Ala Thr Cys Arg
855 860

ACA GTA GCA GTC AAA ATG TTG AAA GAA GGA GCA ACA 2625
Thr Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr
865 870 875

CAC AGT GAG CAT CGA GCT CTC ATG TCT GAA CTC AAG 2661
His Ser Glu His Arg Ala Leu Met Ser Glu Leu Lys
880 885

ATC CTC ATT CAT ATT GGT CAC CAT CTC AAT GTG GTC 2697
Ile Leu Ile His Ile Gly His His Leu Asn Val Val
890 895

```
AAC CTT CTA GGT GCC TGT ACC AAG CCA GGA GGG CCA  2733
Asn Leu Leu Gly Ala Cys Thr Lys Pro Gly Gly Pro
900                 905                 910

CTC ATG GTG ATT GTG GAA TTC TGC AAA TTT GGA AAC  2769
Leu Met Val Ile Val Glu Phe Cys Lys Phe Gly Asn
            915                 920

CTG TCC ACT TAC CTG AGG AGC AAG AGA AAT GAA TTT  2805
Leu Ser Thr Tyr Leu Arg Ser Lys Arg Asn Glu Phe
    925                 930                 935

GTC CCC TAC AAG ACC AAA GGG GCA CGA TTC CGT CAA  2841
Val Pro Tyr Lys Thr Lys Gly Ala Arg Phe Arg Gln
                940                 945

GGG AAA GAC TAC GTT GGA GCA ATC CCT GTG GAT CTG  2877
Gly Lys Asp Tyr Val Gly Ala Ile Pro Val Asp Leu
        950                 955

AAA CGG CGC TTG GAC AGC ATC ACC AGT AGC CAG AGC  2913
Lys Arg Arg Leu Asp Ser Ile Thr Ser Ser Gln Ser
960                 965                 970

TCA GCC AGC TCT GGA TTT GTG GAG GAG AAG TCC CTC  2949
Ser Ala Ser Ser Gly Phe Val Glu Glu Lys Ser Leu
            975                 980

AGT GAT GTA GAA GAA GAG GAA GCT CCT GAA GAT CTG  2985
Ser Asp Val Glu Glu Glu Glu Ala Pro Glu Asp Leu
    985                 990                 995

TAT AAG GAC TTC CTG ACC TTG GAG CAT CTC ATC TGT  3021
Tyr Lys Asp Ph  Leu Thr Leu Glu His Leu Ile Cys
                1000                1005
```

```
TAC AGC TTC CAA GTG GCT AAG GGC ATG GAG TTC TTG   3057
Tyr Ser Phe Gln Val Ala Lys Gly Met Glu Phe Leu
        1010                1015

GCA TCG CGA AAG TGT ATC CAC AGG GAC CTG GCG GCA   3093
Ala Ser Arg Lys Cys Ile His Arg Asp Leu Ala Ala
1020                1025                1030

CGA AAT ATC CTC TTA TCG GAG AAG AAC GTG GTT AAA   3129
Arg Asn Ile Leu Leu Ser Glu Lys Asn Val Val Lys
            1035                1040

ATC TGT GAC TTT GGC TTG GCC CGG GAT ATT TAT AAA   3165
Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys
    1045                1050                1055

GAT CCA GAT TAT GTC AGA AAA GGA GAT GCT CGC CTC   3201
Asp Pro Asp Tyr Val Arg Lys Gly Asp Ala Arg Leu
                1060                1065

CCT TTG AAA TGG ATG GCC CCA GAA ACA ATT TTT GAC   3237
Pro Leu Lys Trp Met Ala Pro Glu Thr Ile Phe Asp
        1070                1075

AGA GTG TAC ACA ATC CAG AGT GAC GTC TGG TCT TTT   3273
Arg Val Tyr Thr Ile Gln Ser Asp Val Trp Ser Phe
1080                1085                1090

GGT GTT TTG CTG TGG GAA ATA TTT TCC TTA GGT GCT   3309
Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Ala
            1095                1100

TCT CCA TAT CCT GGG GTA AAG ATT GAT GAA GAA TTT   3345
Ser Pro Tyr Pro Gly Val Lys Ile Asp Glu Glu Phe
    1105                1110                1115
```

```
TGT AGG CGA TTG AAA GAA GGA ACT AGA ATG AGG GCC   3381
Cys Arg Arg Leu Lys Glu Gly Thr Arg Met Arg Ala
                1120                    1125

CCT GAT TAT ACT ACA CCA GAA ATG TAC CAG ACC ATG   3417
Pro Asp Tyr Thr Thr Pro Glu Met Tyr Gln Thr Met
        1130                    1135

CTG GAC TGC TGG CAC GGG GAG CCC AGT CAG AGA CCC   3453
Leu Asp Cys Trp His Gly Glu Pro Ser Gln Arg Pro
1140                1145                    1150

ACG TTT TCA GAG TTG GTG GAA CAT TTG GGA AAT CTC   3489
Thr Phe Ser Glu Leu Val Glu His Leu Gly Asn Leu
            1155                    1160

TTG CAA GCT AAT GCT CAG CAG GAT GGC AAA GAC TAC   3525
Leu Gln Ala Asn Ala Gln Gln Asp Gly Lys Asp Tyr
    1165                    1170                    1175

ATT GTT CTT CCG ATA TCA GAG ACT TTG AGC ATG GAA   3561
Ile Val Leu Pro Ile Ser Glu Thr Leu Ser Met Glu
                1180                    1185

GAG GAT TCT GGA CTC TCT CTG CCT ACC TCA CCT GTT   3597
Glu Asp Ser Gly Leu Ser Leu Pro Thr Ser Pro Val
        1190                    1195

TCC TGT ATG GAG GAG GAG GAA GTA TGT GAC CCC AAA   3633
Ser Cys Met Glu Glu Glu Glu Val Cys Asp Pro Lys
1200                1205                    1210

TTC CAT TAT GAC AAC ACA GCA GGA ATC AGT CAG TAT   3669
Ph  His Tyr Asp Asn Thr Ala Gly Ile S r Gln Tyr
            1215                    1220
```

```
CTG CAG AAC AGT AAG CGA AAG AGC CGG CCT GTG AGT  3705
Leu Gln Asn Ser Lys Arg Lys Ser Arg Pro Val Ser
    1225                1230                1235

GTA AAA ACA TTT GAA GAT ATC CCG TTA GAA GAA CCA  3741
Val Lys Thr Phe Glu Asp Ile Pro Leu Glu Glu Pro
                1240                1245

GAA GTA AAA GTA ATC CCA GAT GAC AAC CAG ACG GAC  3777
Glu Val Lys Val Ile Pro Asp Asp Asn Gln Thr Asp
        1250                1255

AGT GGT ATG GTT CTT GCC TCA GAA GAG CTG AAA ACT  3813
Ser Gly Met Val Leu Ala Ser Glu Glu Leu Lys Thr
1260                1265                1270

TTG GAA GAC AGA ACC AAA TTA TCT CCA TCT TTT GGT  3849
Leu Glu Asp Arg Thr Lys Leu Ser Pro Ser Phe Gly
            1275                1280

GGA ATG GTG CCC AGC AAA AGC AGG GAG TCT GTG GCA  3885
Gly Met Val Pro Ser Lys Ser Arg Glu Ser Val Ala
    1285                1290                1295

TCT GAA GGC TCA AAC CAG ACA AGC GGC TAC CAG TCC  3921
Ser Glu Gly Ser Asn Gln Thr Ser Gly Tyr Gln Ser
                1300                1305

GGA TAT CAC TCC GAT GAC ACA GAC ACC ACC GTG TAC  3957
Gly Tyr His Ser Asp Asp Thr Asp Thr Thr Val Tyr
        1310                1315

TCC AGT GAG GAA GCA GAA CTT TTA AAG CTG ATA GAG  3993
Ser Ser Glu Glu Ala Glu Leu Leu Lys Leu Il  Glu
1320                1325                1330
```

ATT GGA GTG CAA ACC GGT AGC ACA GCC CAG ATT CTC 4029
Ile Gly Val Gln Thr Gly Ser Thr Ala Gln Ile Leu
1335 1340

CAG CCT GAC ACG GGG ACC ACA CTG AGC TCT CCT CCT 4065
Gln Pro Asp Thr Gly Thr Thr Leu Ser Ser Pro Pro
1345 1350 1355

GTT TAAAAGGAAG CATCCACACC CCAACTCCCG GACATCACAT 4108
Val
1356

GAGAGGTCTG CTCAGATTTT GAAGTGTTGT TCTTTCCACC 4148

AGCAGGAAGT AGCCGCATTT GATTTTCATT TCGACAACAG 4188

AAAAAGGACC TCGGACTGCA GGGAGCCAGC TCTTCTAGGC 4228

TTGTGACC 4236

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 433 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESSS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptid

(ix) FEATURE:

(A) NAME/KEY: ckit proto-oncogene receptor

(B) LOCATION: Amino acids 543-975

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Yarden, Y., et al.

(B) JOURNAL: EMBO J.

(C) VOLUME: 6

(D) PAGES: 3341-3351

(E) DATE: 1987

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
Leu Thr Tyr Lys Tyr Leu Gln Lys Pro Met Tyr Glu Val Gln
543     545                     550                     555

Trp Lys Val Val Glu Glu Ile Asn Gly Asn Asn Tyr Val Tyr
            560                     565                     570

Ile Asp Pro Thr Gln Leu Pro Tyr Asp His Lys Trp Glu Phe
                575                     580

Pro Arg Asn Arg Leu Ser Phe Gly Lys Thr Leu Gly Ala Gly
585                 590                     595

Ala Phe Gly Lys Val Val Ala Glu Thr Ala Tyr Gly Leu Ile
    600                     605                     610

Lys Ser Asp Ala Ala Met Thr Val Ala Val Lys Met Leu Lys
```

615 620 625

Pro Ser Ala His Leu Thr Glu Arg Glu Ala Leu Met Ser Glu
630 635 640

Leu Lys Val Leu Ser Tyr Leu Gly Asn His Met Asn Ile Val
645 650

Asn Leu Leu Gly Ala Cys Thr Ile Gly Gly Pro Thr Leu Val
655 660 665

Ile Thr Glu Tyr Cys Cys Tyr Gly Asp Leu Leu Asn Phe Leu
670 675 680

Arg Arg Lys Arg Asp Ser Phe Ile Cys Ser Lys Gln Glu Asp
685 690 695

His Ala Glu Ala Ala Leu Tyr Lys Asn Leu Leu His Ser Lys
700 705 710

Glu Ser Ser Cys Ser Asp Ser Thr Asn Glu Tyr Met Asp Met
715 720

Lys Pro Gly Val Ser Tyr Val Val Pro Thr Lys Ala Asp Lys
725 730 735

Arg Arg Ser Val Arg Ile Gly Ser Tyr Ile Glu Arg Asp Val
740 745 750

Thr Pro Ala Ile Met Glu Asp Asp Glu Leu Ala Leu Asp Leu
755 760 765

Glu Asp Leu Leu Ser Phe Ser Tyr Gln Val Lys Gly Met Ala
770 775 780

Phe Leu Ala Ser Lys Asn Cys Ile His Arg Asp Leu Ala Ala
785 790

Arg Asn Ile Leu Leu Thr His Gly Arg Ile Thr Lys Ile Cys
795 800 805

Asp Phe Gly Leu Ala Arg Asp Ile Lys Asn Asp Ser Asn Tyr
810 815 820

Val Val Lys Gly Asn Ala Arg Leu Pro Val Lys Val Met Ala
825 830 835

Pro Glu Ser Ile Phe Asn Cys Val Tyr Thr Glu Glu Ser Asp
840 845 850

Val Trp Ser Tyr Gly Ile Phe Leu Trp Glu Leu Phe Ser Leu
855 860

Gly Ser Ser Pro Tyr Pro Gly Met Pro Val Lys Ser Lys Phe
865 870 875

Tyr Lys Met Ile Lys Glu Gly Phe Arg Met Leu Ser Pro Glu
880 885 890

His Ala Pro Ala Glu Met Tyr Asp Ile Met Lys Thr Cys Trp
895 900 905

Asp Ala Asp Pro Leu Lys Arg Pro Thr Phe Lys Gln Ile Val
910 915 920

Gln Leu Ile Glu Lys Gln Ile Ser Glu Ser Thr Asn His Ile
925 930

Tyr Ser Asn Leu Ala Asn Cys Ser Pro Asn Arg Gln Lys Pro
935 940 945

Val Val Asp His Ser Val Arg Ile Asn Ser Val Gly Ser Thr
950 955 960

Ala Ser Ser Ser Gln Pro Leu Leu Val His Asp Asp Val
965 970 975

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 437 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESSS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: CSF-1 receptor

(B) LOCATION: Amino acids 536-972

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Coussens, L., et al.

(B) JOURNAL: Nature

(C) VOLUME: 320

(D) PAGES: 277-280

(E) DATE: 1986

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```
Leu Leu Tyr Lys Tyr Lys Gln Lys Pro Lys Tyr Gln Val Arg
536             540                 545

Trp Lys Ile Ile Glu Ser Tyr Glu Gly Asn Ser Tyr Thr Phe
550                 555                 560

Ile Asp Pro Thr Gln Leu Pro Tyr Asn Glu Lys Trp Glu Phe
    565                 570                 575

Pro Arg Asn Asn Leu Gln Phe Gly Lys Thr Leu Gly Ala Gly
        580                 585                 590

Ala Phe Gly Lys Val Val Glu Ala Thr Ala Phe Gly Leu Gly
            595                 600                 605

Lys Glu Asp Ala Val Leu Lys Val Ala Val Lys Met Leu Lys
                610                 615

Ser Thr Ala His Ala Asp Glu Lys Glu Ala Leu Met Ser Glu
620                 625                 630

Leu Lys Ile Met Ser His Leu Gly Gln His Glu Asn Ile Val
    635                 640                 645

Asn Leu Leu Gly Ala Cys Thr His Gly Gly Pro Val Leu Val
        650                 655                 660

Ile Thr Glu Tyr Cys Cys Tyr Gly Asp Leu Leu Asn Phe Leu
            665                 670                 675
```

Arg Arg Lys Ala Glu Ala Met Leu Gly Pro Ser Leu Ser Pro
680 685

Gly Gln Asp Pro Glu Gly Gly Val Asp Tyr Lys Asn Ile His
690 695 700

Leu Glu Lys Lys Tyr Val Arg Arg Asp Ser Gly Phe Ser Ser
705 710 715

Gln Gly Val Asp Thr Tyr Val Glu Met Arg Pro Val Ser Thr
720 725 730

Ser Ser Asn Asp Ser Phe Ser Glu Gln Asp Leu Asp Lys Glu
735 740 745

Asp Gly Arg Pro Leu Glu Leu Arg Asp Leu Leu His Phe Ser
750 755

Ser Gln Val Ala Gln Gly Met Ala Phe Leu Ala Ser Lys Asn
760 765 770

Cys Ile His Arg Asp Val Ala Ala Arg Asn Val Leu Leu Thr
775 780 785

Asn Gly His Val Ala Lys Ile Gly Asp Phe Gly Leu Ala Arg
790 795 800

Asp Ile Met Asn Asp Ser Asn Tyr Ile Val Lys Gly Asn Ala
805 810 815

Arg Leu Pro Val Lys Trp Met Ala Pro Glu Ser Ile Phe Asp
820 825

Cys Val Tyr Thr Val Gln S r Asp Val Trp Ser Tyr Gly Ile
830 835 840

Leu Leu Trp Glu Ile Phe Ser Leu Gly Leu Asn Pro Tyr Pro
845 850 855

Gly Ile Leu Val Asn Ser Lys Phe Tyr Lys Leu Val Lys Asp
860 865 870

Gly Tyr Gln Met Ala Gln Pro Ala Phe Ala Pro Lys Asn Ile
875 880 885

Tyr Ser Ile Met Gln Ala Cys Trp Ala Leu Glu Pro Thr His
890 895

Arg Pro Thr Phe Gln Gln Ile Cys Ser Phe Leu Gln Glu Gln
900 905 910

Ala Gln Glu Asp Arg Arg Glu Arg Asp Tyr Thr Asn Leu Pro
915 920 925

Ser Ser Ser Arg Ser Gly Gly Ser Gly Ser Ser Ser Ser Glu
930 935 940

Leu Glu Glu Glu Ser Ser Ser Glu His Leu Thr Cys Cys Glu
945 950 955

Gln Gly Asp Ile Ala Gln Pro Leu Leu Gln Pro Asn Asn Tyr
960 965

Gln Phe Cys
970

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 566 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESSS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: PDGF receptor

(B) LOCATION: Amino acids 522-1087

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Gronwald, R., et al.

(B) JOURNAL: Proc. Natl. Acad. Sci.

(C) VOLUME: 85

(D) PAGES: 3435-3439

(E) DATE: 1988

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Met Leu Trp Gln Lys Lys Pro Arg Tyr Glu Ile Arg Trp Lys
522         525                     530                 535

Val Ile Glu Ser Val Ser Ser Asp Gly His Glu Tyr Ile Tyr
                540                 545
```

Val Asp Pro Val Gln Leu Pro Tyr Asp Ser Thr Trp Glu Leu
550 555 560

Pro Arg Asp Gln Leu Val Leu Gly Arg Thr Leu Gly Ser Gly
565 570 575

Ala Phe Gly Gln Val Val Glu Ala Thr Ala His Gly Leu Ser
580 585 590

His Ser Gln Ala Thr Met Lys Val Ala Val Lys Met Leu Lys
595 600 605

Ser Thr Ala Arg Ser Ser Glu Lys Gln Ser Leu Met Ser Glu
610 615

Leu Lys Ile Met Ser His Leu Gly Pro His Leu Asn Val Val
620 625 630

Asn Leu Leu Gly Ala Cys Thr Lys Gly Gly Pro Ile Tyr Ile
635 640 645

Ile Thr Glu Tyr Cys Arg Tyr Gly Asp Leu Val Asp Tyr Leu
650 655 660

His Arg Asn Lys His Thr Phe Leu Gln Arg His Ser Asn Lys
665 670 675

His Cys Pro Pro Ser Ala Glu Leu Tyr Ser Asn Ala Leu Pro
680 685

Val Gly Phe Ser Leu Pro Ser His Leu Asn Leu Thr Gly Glu
690 695 700

Ser Asp Gly Gly Tyr M t Asp Met Ser Lys Asp Glu S r Ile
705 710 715

Asp Tyr Val Pro Met Leu Asp Met Lys Gly Asp Ile Lys Tyr
720 725 730

Ala Asp Ile Glu Ser Pro Ser Tyr Met Ala Pro Tyr Asp Asn
735 740 745

Tyr Val Pro Ser Ala Pro Glu Arg Thr Tyr Arg Ala Thr Leu
750 755

Ile Asn Asp Ser Pro Val Leu Ser Tyr Thr Asp Leu Val Gly
760 765 770

Phe Ser Tyr Gln Val Ala Asn Gly Met Asp Phe Leu Ala Ser
775 780 785

Lys Asn Cys Val His Arg Asp Leu Ala Ala Arg Asn Val Leu
790 795 800

Ile Cys Glu Gly Lys Leu Val Lys Ile Cys Asp Phe Gly Phe
805 810 815

Ala Arg Asp Ile Met Arg Asp Ser Asn Tyr Ile Ser Lys Gly
820 825

Ser Thr Tyr Leu Pro Leu Lys Trp Met Ala Pro Glu Ser Ile
830 835 840

Phe Asn Ser Leu Tyr Thr Thr Leu Ser Asp Val Trp Ser Phe
845 850 855

Gly Ile Leu Leu Trp Glu Ile Phe Thr Leu Gly Gly Thr Pro
860 865 870

Tyr Pro Glu Leu Pro Met Asn Asp Gln Phe Tyr Asn Ala Ile
875 880 885

Lys Arg Gly Tyr Arg Met Ala Gln Pro Ala His Ala Ser Asp
890 895

Glu Ile Tyr Glu Ile Met Gln Lys Cys Trp Glu Glu Lys Phe
900 905 910

Glu Thr Arg Pro Pro Phe Ser Gln Leu Val Leu Leu Leu Glu
915 920 925

Arg Leu Leu Gly Glu Gly Tyr Lys Lys Lys Tyr Gln Gln Val
930 935 940

Asp Glu Glu Phe Leu Arg Ser Asp His Pro Ala Ile Leu Arg
945 950 955

Ser Gln Ala Arg Phe Pro Gly Ile His Ser Leu Arg Ser Pro
960 965

Leu Asp Thr Ser Ser Val Leu Tyr Thr Ala Val Gln Pro Asn
970 975 980

Glu Ser Asp Asn Asp Tyr Ile Ile Pro Leu Pro Asp Pro Lys
985 990 995

Pro Asp Val Ala Asp Glu Gly Leu Pro Glu Gly Ser Pro Ser
1000 1005 1010

Leu Ala Ser Ser Thr Leu Asn Glu Val Asn Thr Ser Ser Thr
1015 1020 1025

Ile Ser Cys Asp Ser Pro Leu Glu Leu Gln Glu Glu Pro Gln
1030 1035

Gln Ala Glu Pro Glu Ala Gln Leu Glu Gln Pro Gln Asp Ser
1040 1045 1050

Gly Cys Pro Gly Pro Leu Ala Glu Ala Glu Asp Ser Phe Leu
1055 1060 1065

Glu Gln Pro Gln Asp Ser Gly Cys Pro Gly Pro Leu Ala Glu
1070 1075 1080

Ala Glu Asp Ser Phe Leu
1085

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TCGACGCGCG ATG GAG 16

We claim:

1. An isolated DNA sequence encoding the Kinase insert Domain containing Receptor.

2. The DNA sequence of Claim 1 wherein said sequence is a human gene.

3. An isolated DNA sequence comprising a DNA sequence capable of hybridizing with a DNA sequence encoding the amino acid sequence of Figure 7.

4. A method for the production of a growth factor receptor which comprises transforming a host cell with the DNA sequence of Claim 3 and culturing the host cell under conditions which result in expression of the gene by an expression vector.

5. The method of Claim 4 wherein the host cell is a bacteria, virus, yeast, insect or mammalian cell line.

6. The method of Claim 5 wherein the host cell is a COS-1 cell, NIH3T3 fibroblast or CMT-3 monkey kidney cell.

7. The method of Claim 5 where the expression vector is pcDNA1tkpASP expression vector.

8. A lambda gt11 phage harboring the clone BTIII081.8 (ATCC accession number 40,931) or the clone BTIII129.5 (ATCC Accession number 40,975).

9. A plasmid pBlueScript KS which contains the clone BTIV169 (ATCC accession number 75200).

10. An isolated growth factor receptor designated the Kinase insert Domain containing Receptor.

11. The receptor of Claim 10 comprising the amino acid sequence of Figure 7.

12. The receptor of Claim 10 encoded by an isolat d DNA s quence comprising a DNA sequenc capable

of hybridizing with a DNA sequence encoding the amino acid sequence of Figure 7.

13. A biologically active protein fragment which retains the receptor activity of the receptor of Claim 10.

14. An isolated DNA sequence encoding a biologically active protein fragment which retains the receptor activity of an isolated growth factor receptor designated the Kinase insert Domain containing Receptor.

15. An oligonucleotide primer consisting of an oligonucleotide primer having 21 bases and having a sequence depicted for Primer 1 in Figure 2.

16. An oligonucleotide primer consisting of an oligonucleotide primer having 29 bases and having a sequence depicted for Primer 2 in Figure 2.

17. The 363 base pair product having the sequence depicted in Figure 4, or a biological equivalent of said sequence.

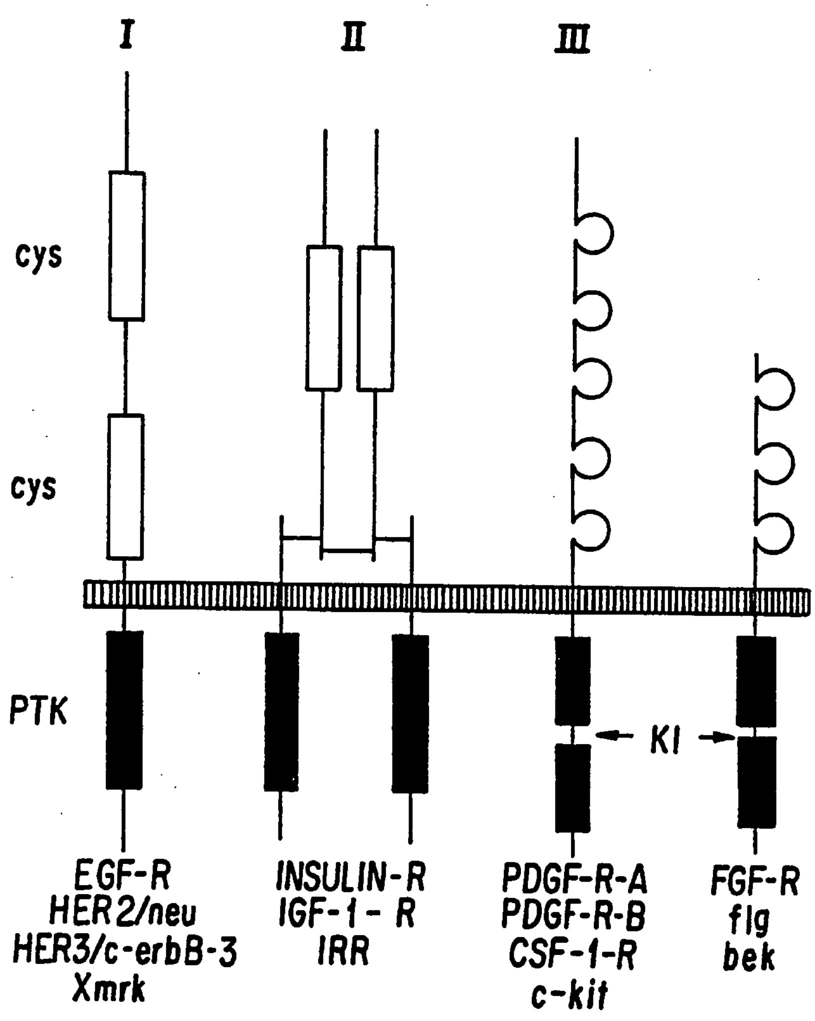

I
II
III
cys
cys
PTK
KI
EGF-R
HER2/neu
HER3/c-erbB-3
Xmrk
INSULIN-R
IGF-1- R
IRR
PDGF-R-A
PDGF-R-B
CSF-1-R
c-kit
FGF-R
flg
bek

FIG. 1

```
PRIMER 1
RECEPTOR

  PDGF             AAC CTG TTG GGG GCC TGC ACC
  ckit              T   A   T   A
  CSF               T           A
  FGF                       C                 G
------------------------------------------------------------
PRIMER 1    GTCGAC AAC CTG TTG GGG GCC TGC AAC
                    T           A

PRIMER 2
RECEPTOR

  PDGF             CAC AGA GAC CTG GCG GCT AGG AAC GTG CT
  ckit                         T    GA   C   A   T     A
  CSF                  C G     G     A  GC C T
  FGF                  C             C   C       T   C
------------------------------------------------------------
CONSENSUS          CAC AGA GAC CTG GCC GCT AGI AAC GTG CT
                       C             T   C       T

PRIMER 2    GAATTC AG CAC GTT ICT AGC CGC CAG GTC TCT GTG
                          T       G   T             G
```

FIG.2

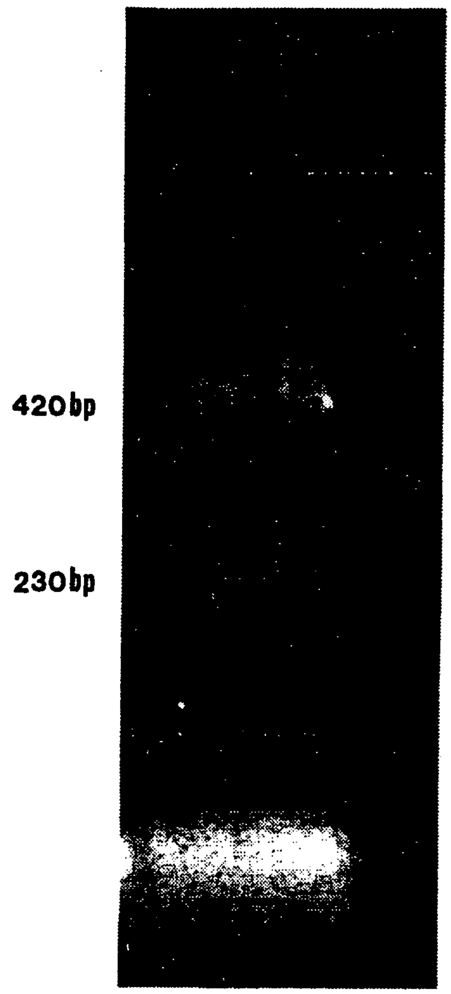


FIG. 3

```
GAATTCTGCAAATTTGGAAACCTGTCCACTT
ACCTGAGGACGAAGAGAAATGAATTTGTCCC
CTACAAGACCAAAGGGGCACGATTCCGTCAA
GGGAAAGACTACGTTGGAGCAATCCCTGTGG
ATCTGAAACGGCGCTTGGACACGCATCACCA
GTAGCCAGAGCTCAGCCAGCTCTGGATTTGT
GGAGGAGAAGTCCCTCAGTGATGTAGAAGAA
GAGGAAGCTCCTGAAGATCTGTATAAGGACT
TCCTGACCTTGGAGCATCTCATCTGTTACAG
TTTCCAAGTGGCTAAGGGCATGGAGTTCTTG
GCATCGCGAAAGTGTATCCACAGAGACCTGG
CAGCCAGGAACGTGCTGAATTC
```

FIG.4A

```
GTCGACAATCTGTTGGGGGCCTGCACCATCC
CAACATCCTGCTGCTCTACAACTATTTTTAT
GACCGGAGGAGGATCTACTTGATTCTAGAGT
ATGCCCCCCGCGGAGCTCTACAAGGAGCTGC
AGAAGAGCTGCACATTTGACGAGCAGCGAAC
AGCCACGATCATGGAGGAGTTGGCAGATGCT
CTAATGTACTGCCGTGGGAAGAAGGTGATTC
ACAGAGACCTGGCAGCCAGCAACGTGCTGAA
TTC
```

FIG.4B

```
               10v       20v       30v       40v       50v       60v       70v       80v       90v       100v
PDGF    AACCTGTGGGGGCCTGCACCAAAGGAGGACCATCTATATCATCTATATCATCACTGAGTACTGCCGCTACGGAGACCTGGTGGACTACCTGCACCGCAACA
                                                                   TGC   T  GGA ACCTG       TACCTG   GCAA A
360 bp                                                             TGCAAATTTGGAAACCTGTCCACTTACCTGAGGAGCAAGA
                                                                          10^       20^       30^       40^

                         110v      120v      130v      140v      150v      160v      170v
PDGF    A-----------------ACAC-ACCTTCCTGCAGCACCACTCCGACAAGCGCCGCCCGCCCAGCGCGGAGCT-CTAC-----AGCAATGCTCTGCCCGT
                          ACA  ACC    G  GCAC A TCCG CAAG                GGA    CTAC     AGCAAT      CCC
360 bp  GAAATGAATTTGTCCCCTACAAGACCAAAGGG--GCACGATTCCGTCAAG-----------------GGAAAGACTACGTTGGAGCAAT-------CCCTG
           50^       60^       70^       80^                            90^       100^      110^

         v     190v      200v      210v      220v      230v      240v      250v      260v      270v
PDGF    TGGGCTCCCCCTGCCCAGCCATGTGTCCTTGACCGGGGGAGAGCGACGGTGGCTACATGGACATGAGCAAGGACGAGTCGGTGGACTATGTGCCCATGCT
        TGG        G C       G          G  AGAGC   G   GCT   TGGA  TG G  AGGA  AGTC  T     ATGT
360 bp  TGGATCTGAAACGGCGCTTGGACAGCATCACCAGTAGCCAGAGCTCAGCCAGCTC--TGGATTTGTGG-AGGAGAAGTCCCTCAGTGATGTAGAAGAAGA
         120^      130^      140^      150^      160^      170^      180^      190^      200^      210^

         v     290v      300v      310v      320v      330v      340v      350v      360v      370v
PDGF    GGACATGAAAGGAGACGTCAAATAGCAGACATCGAGTCCTCCAACTACATGGCCCCTTACGATAACTACGTTCCCTCTGCCCCTGAGAGGACCTGCCGAG
        GGA       G AGA
                                                                                             CTG   A
        GGAAGCTCCTGAAGAT-----------------------------------------------------------------------CTGTATAA
360 bp     220^                                                                                230^

         v     390v      400v      410v      420V      430V      440v      450v      460v      470v
PDGF    CAACTTTGATCAACGAGTCTCCAGTGCTAAGCTACATGGAC---CTCGTGGGCTTCAGCTACCAGGTGGCCAATGGCATGGAGTTCTGGCCTCCAAGAAC
          ACTT                                 TGGA   CTC T  G T CAGCT CCA GTGG  A  GGCATGGAGTTCT G C  C  GAA
360 bp  GGACTTCCTGACCT------------------------TGGAGCATCTCATCTGTTACAGCTTCCAAGTGGTAAG-GGCATGGAGTTCTTGGCATCGCGAAA
        240^      250^                             260^      270^      280^      290^      300^      310^

           480v      490v      500v      510v
PDGF    TGCG-TCCACAGAGACCTGGCGGCTAGGAACGTGCTT
             TCCACAGAGACCTGGC GC AGGAACGTGCT
360 bp  GTGTATCCACAGAGACCTGGCAGCCAGGAACGTGCT
          320^      330^      340^
```

FIG.5A

```
              10v       20v       30v       40V       50v       60v       70v       80v
FGF     AACCTGCTGGGGGCCTGCACGCAGGATGGTCCCTTGTATGTCATCGTGGAGTATGCCTCCAAGGGCAACCTGCGGGAGTACCTGC----------------
        AA CTG TGGGGGCCTGCAC          CCT          CGT         CTC A   CAAC       A  ACC G
230 bp  AATCTGTTGGGGGCCTGCACCATCCCAACATCCTG----------CGT----------CTCTA----CAACTATTTTTATGACCGGAGGAGGATCTACTTGAT
              10^       20^       30^                          40^       50^       60^       70^
           90v       100v      110v      120v      130v      140v      150v      160v      170v      180v
FGFI     ---AGACCCGGAGGCCCCCAGGGCTGGAATACTGCTATAACCCCAGCCACAACCCAGAGGAGCAGCTCTCCTCCAAGGACCTGGTGTCCTGCGCCTACCA
            AGA    G   CCC C G GCT  A  A        A    AGC  CA    GA GAGCAGC   C  CCA G  C TGG                A
230 bp   TCTAGAGTATGCCCCCCGCGGAGCTCTACAAGGAGCTGCAGAAGAGCTGCACATTTGACGAGCAGCGAACAGCCACGATCATGG----------------A
         ^       90^       100^      110^      120^      130^      140^      150^      160^
             190v      200v      210v      220v      230v      240v      250v
FGF      GGA---GGCCCGAGGCATGGAGTATCTGGCCTCCAAGAAGTGCATACACCGAGACCTGGCAGCCAGGAATGTCCTG
         GGA   GGC    G   T   GTA         AAGAAG   AT CAC GAGACCTGGCAGCCAG AA GT CT
230 bp   GGAGTTGGCAGATGCTCTAATGTACTGCCGTGGGAAGAAGGTGATTCACAGAGACCTGGCAGCCAGCAACGTGCT
             170^      180^      190^      200^      210^      220^      230^
```

FIG.5B

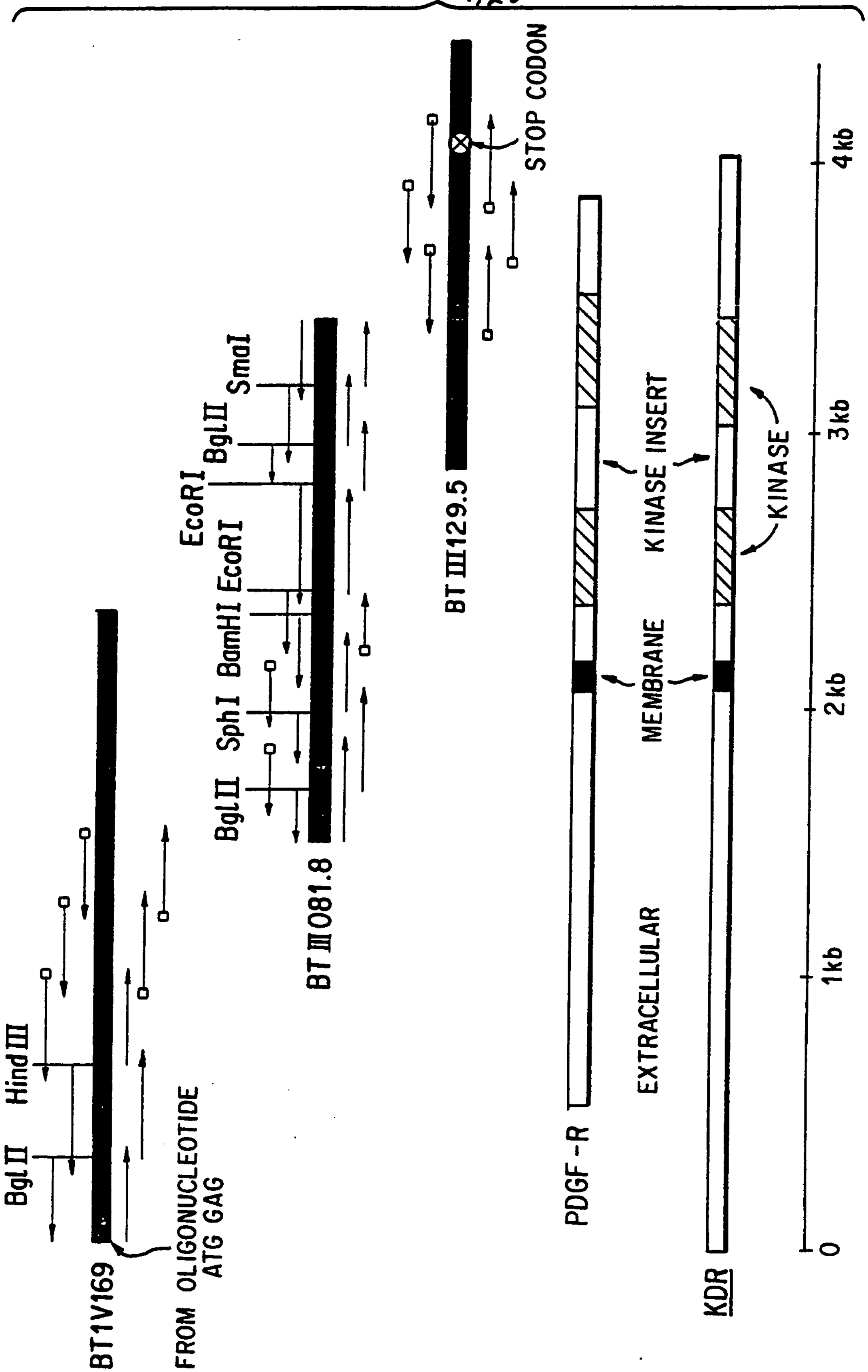

BT1V169
FROM OLIGONUCLEOTIDE
ATG GAG
BglII
HindIII
BT III 081.8
BglII
SphI
BamHI
EcoRI
EcoRI
BglII
SmaI
BT III 129.5
STOP CODON
PDGF-R
KDR
EXTRACELLULAR
MEMBRANE
KINASE INSERT
KINASE
0
1kb
2kb
3kb
4kb

FIG. 6

```
                         10                  20                  30                  40                  50
                          *                   *                   *                   *                   *
ATG GAG AGC AAG GTG CTG CTG GCC GTC GCC CTG TGG CTC TGC GTG GAG ACC CGG
Met Glu Ser Lys Val Leu Leu Ala Val Ala Leu Trp Leu [Cys] Val Glu Thr Arg>

                60                  70                  80                  90                 100
                 *                   *                   *                   *                   *
GCC GCC TCT GTG GGT TTG CCT AGT GTT TCT CTT GAT CTG CCC AGG CTC AGC ATA
Ala Ala Ser Val Gly Leu Pro Ser Val Ser Leu Asp Leu Pro Arg Leu Ser Ile>

110                 120                 130                 140                 150                 160
  *                   *                   *             *     *                   *                   *
CAA AAA GAC ATA CTT ACA ATT AAG GCT AAT ACA ACT CTT CAA ATT ACT TGC AGG
Gln Lys Asp Ile Leu Thr Ile Lys Ala Asn Thr Thr Leu Gln Ile Thr [Cys] Arg>

            170                 180                 190                 200                 210
              *                   *                   *             *     *                   *
GGA CAG AGG GAC TTG GAC TGG CTT TGG CCC AAT AAT CAG AGT GGC AGT GAG CAA
Gly Gln Arg Asp Leu Asp Trp Leu Trp Pro Asn Asn Gln Ser Gly Ser Glu Gln>

      220                 230                 240                 250                 260                 270
        *                   *                   *                   *                   *                   *
AGG GTG GAG GTG ACT GAG TGC AGC GAT GGC CTC TTC TGT AAG ACA CTC ACA ATT
Arg Val Glu Val Thr Glu [Cys] Ser Asp Gly Leu Phe [Cys] Lys Thr Leu Thr Ile>

                280                 290                 300                 310                 320
                  *             *     *                   *                   *                   *
CCA AAA GTG ATC GGA AAT GAC ACT GGA GCC TAC AAG TGC TTC TAC CGG GAA ACT
Pro Lys Val Ile Gly Asn Asp Thr Gly Ala Tyr Lys [Cys] Phe Tyr Arg Glu Thr>
```

FIG. 7A

```
                330                 340                 350                 360                 370
                  *                   *                   *                   *                   *
GAC TTG GCC TCG GTC ATT TAT GTC TAT GTT CAA GAT TAC AGA TCT CCA TTT ATT
Asp Leu Ala Ser Val Ile Tyr Val Tyr Val Gln Asp Tyr Arg Ser Pro Phe Ile>

380                 390                 400                 410                 420               430
  *                   *                   *                   *                   *             *   *
GCT TCT GTT AGT GAC CAA CAT GGA GTC GTG TAC ATT ACT GAG AAC AAA AAC AAA
Ala Ser Val Ser Asp Gln His Gly Val Val Tyr Ile Thr Glu Asn Lys Asn Lys>

          440                 450                 460                 470                 480
            *                   *                   *                   *     *             *
ACT GTG GTG ATT CCA TGT CTC GGG TCC ATT TCA AAT CTC AAC GTG TCA CTT TGT
Thr Val Val Ile Pro Cys Leu Gly Ser Ile Ser Asn Leu Asn Val Ser Leu Cys>

      490                 500                 510                 520                 530                 540
        *                   *                   *                   *                   *                   *
GCA AGA TAC CCA GAA AAG AGA TTT GTT CCT GAT GGT AAC AGA ATT TCC TGG GAC
Ala Arg Tyr Pro Glu Lys Arg Phe Val Pro Asp Gly Asn Arg Ile Ser Trp Asp>

              550                 560                 570                 580                 590
                *                   *                   *                   *                   *
AGC AAG AAG GGC TTT ACT ATT CCC AGC TAC ATG ATC AGC TAT GCT GGC ATG GTC
Ser Lys Lys Gly Phe Thr Ile Pro Ser Tyr Met Ile Ser Tyr Ala Gly Met Val>

        600                 610                 620                 630                 640
          *                   *                   *                   *                   *
TTC TGT GAA GCA AAA ATT AAT GAT GAA AGT TAC CAG TCT ATT ATG TAC ATA GTT
Phe Cys Glu Ala Lys Ile Asn Asp Glu Ser Tyr Gln Ser Ile Met Tyr Ile Val>
```

FIG. 7B

10/28

```
650                 660                     670                   680                   690                   700
 *                   *                       *                     *                     *                     *
GTC GTT GTA GGG TAT AGG ATT TAT GAT GTG GTT CTG AGT CCG TCT CAT GGA ATT
Val Val Val Gly Tyr Arg Ile Tyr Asp Val Val Leu Ser Pro Ser His Gly Ile>

        710                 720                   730       *       740                   750
         *                   *                     *                 *                     *
GAA CTA TCT GTT GGA GAA AAG CTT GTC TTA AAT TGT ACA GCA AGA ACT GAA CTA
Glu Leu Ser Val Gly Glu Lys Leu Val Leu Asn Cys Thr Ala Arg Thr Glu Leu>

   760                 770                   780                   790                   800                   810
    *                   *                     *                     *                     *                     *
AAT GTG GGG ATT GAC TTC AAC TGG GAA TAC CCT TCT TCG AAG CAT CAG CAT AAG
Asn Val Gly Ile Asp Phe Asn Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys>

            820                 830                   840                   850                   860
             *                   *                     *                     *                     *
AAA CTT GTA AAC CGA GAC CTA AAA ACC CAG TCT GGG AGT GAG ATG AAG AAA TTT
Lys Leu Val Asn Arg Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe>

       870                   880                   890                   900                   910
        *                     *                     *                     *                     *
TTG AGC ACC TTA ACT ATA GAT GGT GTA ACC CGG AGT GAC CAA GGA TTG TAC ACC
Leu Ser Thr Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu Tyr Thr>

920                 930                   940                   950       *       960                   970
 *                   *                     *                     *                 *                     *
TGT GCA GCA TCC AGT GGG CTG ATG ACC AAG AAG AAC AGC ACA TTT GTC AGG GTC
Cys Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val Arg Val>
```

FIG. 7C

11/28

```
           980                 990                 1000                1010                1020
            *                   *                   *                   *                   *
CAT GAA AAA CCT TTT GTT GCT TTT GGA AGT GGC ATG GAA TCT CTG GTG GAA GCC
His Glu Lys Pro Phe Val Ala Phe Gly Ser Gly Met Glu Ser Leu Val Glu Ala>

  1030                1040                1050                1060                1070                1080
    *                   *                   *                   *                   *                   *
ACG GTG GGG GAG CGT GTC AGA ATC CCT GCG AAG TAC CTT GGT TAC CCA CCC CCA
Thr Val Gly Glu Arg Val Arg Ile Pro Ala Lys Tyr Leu Gly Tyr Pro Pro Pro>

             1090                1100                1110                1120                1130
               *                   *                   *                   *                   *
GAA ATA AAA TGG TAT AAA AAT GGA ATA CCC CTT GAG TCC AAT CAC ACA ATT AAA
Glu Ile Lys Trp Tyr Lys Asn Gly Ile Pro Leu Glu Ser Asn His Thr Ile Lys>

      1140                1150                1160                1170                1180
        *                   *                   *                   *                   *     *
GCG GGG CAT GTA CTG ACG ATT ATG GAA GTG AGT GAA AGA GAC ACA GGA AAT TAC
Ala Gly His Val Leu Thr Ile Met Glu Val Ser Glu Arg Asp Thr Gly Asn Tyr>

1190                1200                1210                1220                1230                1240
   *                   *                   *                   *                   *                   *
ACT GTC ATC CTT ACC AAT CCC ATT TCA AAG GAG AAG CAG AGC CAT GTG GTC TCT
Thr Val Ile Leu Thr Asn Pro Ile Ser Lys Glu Lys Gln Ser His Val Val Ser>

        1250                1260                1270                1280                1290
          *                   *                   *                   *                   *
CTG GTT GTG TAT GTC CCA CCC CAG ATT GGT GAG AAA TCT CTA ATC TCT CCT GTG
Leu Val Val Tyr Val Pro Pro Gln Ile Gly Glu Lys Ser Leu Ile Ser Pro Val>
```

FIG. 7D

SUBSTITUTE SHEET

```
1300        1310        1320        1330        1340        1350
   *           *           *           *           *           *
GAT TCC TAC CAG TAC GGC ACC ACT CAA ACG CTG ACA TGT ACG GTC TAT GCC ATT
Asp Ser Tyr Gln Tyr Gly Thr Thr Gln Thr Leu Thr[Cys]Thr Val Tyr Ala Ile>

              1360        1370        1380        1390        1400
                 *           *           *           *           *
CCT CCC CCG CAT CAC ATC CAC TGG TAT TGG CAG TTG GAG GAA GAG TGC GCC AAC
Pro Pro Pro His His Ile His Trp Tyr Trp Gln Leu Glu Glu Glu[Cys]Ala Asn>

      1410        1420        1430        1440        1450
         *           *           *           *           *
GAG CCC AGC CAA GCT GTC TCA GTG ACA AAC CCA TAC CCT TGT GAA GAA TGG AGA
Glu Pro Ser Gln Ala Val Ser Val Thr Asn Pro Tyr Pro[Cys]Glu Glu Trp Arg>

1460        1470        1480        1490        1500        1510
   *           *           *           *           *           *
AGT GTG GAG GAC TTC CAG GGA GGA AAT AAA ATT GAA GTT AAT AAA AAT CAA TTT
Ser Val Glu Asp Phe Gln Gly Gly Asn Lys Ile Glu Val Asn Lys Asn Gln Phe>

        1520        1530        1540        1550        1560
           *           * *         *           *           *
GCT CTA ATT GAA GGA AAA AAC AAA ACT GTA AGT ACC CTT GTT ATC CAA GCG GCA
Ala Leu Ile Glu Gly Lys Asn Lys Thr Val Ser Thr Leu Val Ile Gln Ala Ala>

1570        1580        1590        1600        1610        1620
* *           *           *           *           *           *
AAT GTG TCA GCT TTG TAC AAA TGT GAA GCG GTC AAC AAA GTC GGG AGA GGA GAG
Asn Val Ser Ala Leu Tyr Lys[Cys]Glu Ala Val Asn Lys Val Gly Arg Gly Glu>
```

13/28

FIG. 7E

13/28

```
          1630       1640       1650       1660       1670
             *          *          *          *          *
AGG GTG ATC TCC TTC CAC GTG ACC AGG GGT CCT GAA ATT ACT TTG CAA CCT GAC
Arg Val Ile Ser Phe His Val Thr Arg Gly Pro Glu Ile Thr Leu Gln Pro Asp>

    1680       1690       1700       1710       1720
       *          *          *          *          *
ATG CAG CCC ACT GAG CAG GAG AGC GTG TCT TTG TGG TGC ACT GCA GAC AGA TCT
Met Gln Pro Thr Glu Gln Glu Ser Val Ser Leu Trp [Cys] Thr Ala Asp Arg Ser>

1730       1740       1750       1760       1770       1780
   *          *          *          *          *          *
ACG TTT GAG AAC CTC ACA TGG TAC AAG CTT GGC CCA CAG CCT CTG CCA ATC CAT
Thr Phe Glu Asn Leu Thr Trp Tyr Lys Leu Gly Pro Gln Pro Leu Pro Ile His>

      1790       1800       1810       1820       1830
         *          *          *          *          *
GTG GGA GAG TTG CCC ACA CCT GTT TGC AAG AAC TTG GAT ACT CTT TGG AAA TTG
Val Gly Glu Leu Pro Thr Pro Val [Cys] Lys Asn Leu Asp Thr Leu Trp Lys Leu>

  1840       1850       1860       1870       1880       1890
 *   *          *      *   *          *          *          *
AAT GCC ACC ATG TTC TCT AAT AGC ACA AAT GAC ATT TTG ATC ATG GAG CTT AAG
Asn Ala Thr Met Phe Ser Asn Ser Thr Asn Asp Ile Leu Ile Met Glu Leu Lys>

          1900       1910       1920       1930       1940
 *           *          *          *          *          *
AAT GCA TCC TTG CAG GAC CAA GGA GAC TAT GTC TGC CTT GCT CAA GAC AGG AAG
Asn Ala Ser Leu Gln Asp Gln Gly Asp Tyr Val Cys Leu Ala Gln Asp Arg Lys>
```

FIG. 7F

```
       1950                1960                1970                1980                1990
          *                   *                   *                   *                   *
ACC AAG AAA AGA CAT TGC GTG GTC AGG CAG CTC ACA GTC CTA GAG CGT GTG GCA
Thr Lys Lys Arg His Cys Val Val Arg Gln Leu Thr Val Leu Glu Arg Val Ala>

2000                2010                2020                2030                2040                2050
   *                   *                   *   *               *                   *                   *
CCC ACG ATC ACA GGA AAC CTG GAG AAT CAG ACG ACA AGT ATT GGG GAA AGC ATC
Pro Thr Ile Thr Gly Asn Leu Glu Asn Gln Thr Thr Ser Ile Gly Glu Ser Ile>

           2060                2070                2080                2090                2100
              *                   *                   *                   *                   *
GAA GTC TCA TGC ACG GCA TCT GGG AAT CCC CCT CCA CAG ATC ATG TGG TTT AAA
Glu Val Ser Cys Thr Ala Ser Gly Asn Pro Pro Pro Gln Ile Met Trp Phe Lys>

   2110                2120                2130                2140                2150                2160
      **                  *                   *                   *                   *                   *
GAT AAT GAG ACC CTT GTA GAA GAC TCA GGC ATT GTA TTG AAG GAT GGG AAC CGG
Asp Asn Glu Thr Leu Val Glu Asp Ser Gly Ile Val Leu Lys Asp Gly Asn Arg>

               2170                2180                2190                2200                2210
 *                *                   *                   *                   *                   *
AAC CTC ACT ATC CGC AGA GTG AGG AAG GAG GAC GAA GGC CTC TAC ACC TGC CAG
Asn Leu Thr Ile Arg Arg Val Arg Lys Glu Asp Glu Gly Leu Tyr Thr Cys Gln>

       2220                2230                2240                2250                2260
          *                   *                   *                   *                   *
GCA TGC AGT GTT CTT GGC TGT GCA AAA GTG GAG GCA TTT TTC ATA ATA GAA GGT
Ala Cys Ser Val Leu Gly Cys Ala Lys Val Glu Ala Phe Phe Ile Ile Glu Gly>
```

FIG. 7G

15/28

```
2270       2280       2290       2300       2310       2320
   *          *          *          *          *          *
GCC CAG GAA AAG ACG AAC TTG GAA ATC ATT ATT CTA GTA GGC ACG ACG GTG ATT
Ala Gln Glu Lys Thr Asn Leu Glu Ile Ile Ile Leu Val Gly Thr Thr Val Ile>

      2330       2340       2350       2360       2370
         *          *          *          *          *
GCC ATG TTC TTC TGG CTA CTT CTT GTC ATC ATC CTA GGG ACC GTT AAG CGG GCC
Ala Met Phe Phe Trp Leu Leu Leu Val Ile Ile Leu Gly Thr Val Lys Arg Ala>

 2380       2390       2400       2410       2420       2430
    *          *          *          *          *          *
AAT GGA GGG GAA CTG AAG ACA GGC TAC TTG TCC ATC GTC ATG GAT CCA GAT GAA
Asn Gly Gly Glu Leu Lys Thr Gly Tyr Leu Ser Ile Val Met Asp Pro Asp Glu>

        2440       2450       2460       2470       2480
           *          *          *          *          *
CTC CCA TTG GAT GAA CAT TGT GAA CGA CTG CCT TAT GAT GCC AGC AAA TGG GAA
Leu Pro Leu Asp Glu His Cys Glu Arg Leu Pro Tyr Asp Ala Ser Lys Trp Glu>

   2490       2500       2510       2520       2530
      *          *          *          *          *
TTC CCC AGA GAC CGG CTG AAC CTA GGT AAG CCT CTT GGC CGT GGT GCC TTT GGC
Phe Pro Arg Asp Arg Leu Asn Leu Gly Lys Pro Leu Gly Arg Gly Ala Phe Gly>

2540       2550       2560       2570       2580       2590
   *          *          *          *          *          *
CAA GAG ATT GAA GCA GAT GCC TTT GGA ATT GAC AAG ACA GCA ACT TGC AGG ACA
Gln Glu Ile Glu Ala Asp Ala Phe Gly Ile Asp Lys Thr Ala Thr Cys Arg Thr>
```

FIG. 7H

```
     2600      2610       2620       2630      2640
        *         *          *          *         *
GTA GCA GTC AAA ATG TTG AAA GAA GGA GCA ACA CAC AGT GAG CAT CGA GCT CTC
Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr His Ser Glu His Arg Ala Leu>

 2650       2660       2670       2680       2690       2700
     *          *          *          *          *         *
ATG TCT GAA CTC AAG ATC CTC ATT CAT ATT GGT CAC CAT CTC AAT GTG GTC AAC
Met Ser Glu Leu Lys Ile Leu Ile His Ile Gly His His Leu Asn Val Val Asn>

        2710       2720       2730       2740       2750
           *          *          *          *          *
CTT CTA GGT GCC TGT ACC AAG CCA GGA GGG CCA CTC ATG GTG ATT GTG GAA TTC
Leu Leu Gly Ala Cys Thr Lys Pro Gly Gly Pro Leu Met Val Ile Val Glu Phe>

    2760       2770       2780       2790       2800
       *          *          *          *          *
TGC AAA TTT GGA AAC CTG TCC ACT TAC CTG AGG AGC AAG AGA AAT GAA TTT GTC
Cys Lys Phe Gly Asn Leu Ser Thr Tyr Leu Arg Ser Lys Arg Asn Glu Phe Val>

2810       2820       2830       2840       2850       2860
   *          *          *          *          *          *
CCC TAC AAG ACC AAA GGG GCA CGA TTC CGT CAA GGG AAA GAC TAC GTT GGA GCA
Pro Tyr Lys Thr Lys Gly Ala Arg Phe Arg Gln Gly Lys Asp Tyr Val Gly Ala>

      2870       2880       2890       2900       2910
         *          *          *          *          *
ATC CCT GTG GAT CTG AAA CGG CGC TTG GAC AGC ATC ACC AGT AGC CAG AGC TCA
Ile Pro Val Asp Leu Lys Arg Arg Leu Asp Ser Ile Thr Ser Ser Gln Ser Ser>
```

FIG. 7I

2920 * 2930 * 2940 * 2950 * 2960 * 2970 *
GCC AGC TCT GGA TTT GTG GAG GAG AAG TCC CTC AGT GAT GTA GAA GAA GAG GAA
Ala Ser Ser Gly Phe Val Glu Glu Lys Ser Leu Ser Asp Val Glu Glu Glu Glu>

2980 * 2990 * 3000 * 3010 * 3020 *
GCT CCT GAA GAT CTG TAT AAG GAC TTC CTG ACC TTG GAG CAT CTC ATC TGT TAC
Ala Pro Glu Asp Leu Tyr Lys Asp Phe Leu Thr Leu Glu His Leu Ile Cys Tyr>

3030 * 3040 * 3050 * 3060 * 3070 *
AGC TTC CAA GTG GCT AAG GGC ATG GAG TTC TTG GCA TCG CGA AAG TGT ATC CAC
Ser Phe Gln Val Ala Lys Gly Met Glu Phe Leu Ala Ser Arg Lys Cys Ile His>

3080 * 3090 * 3100 * 3110 * 3120 * 3130 *
AGG GAC CTG GCG GCA CGA AAT ATC CTC TTA TCG GAG AAG AAC GTG GTT AAA ATC
Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser Glu Lys Asn Val Val Lys Ile>

3140 * 3150 * 3160 * 3170 * 3180 *
TGT GAC TTT GGC TTG GCC CGG GAT ATT TAT AAA GAT CCA GAT TAT GTC AGA AAA
Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys Asp Pro Asp Tyr Val Arg Lys>

3190 * 3200 * 3210 * 3220 * 3230 * 3240 *
GGA GAT GCT CGC CTC CCT TTG AAA TGG ATG GCC CCA GAA ACA ATT TTT GAC AGA
Gly Asp Ala Arg Leu Pro Leu Lys Trp Met Ala Pro Glu Thr Ile Phe Asp Arg>

FIG. 7J

15/25

```
              3250                3260                3270                  3280                3290
                *                   *                   *                     *                   *
GTG TAC ACA ATC CAG AGT GAC GTC TGG TCT TTT GGT GTT TTG CTG TGG GAA ATA
Val Tyr Thr Ile Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile>

        3300                3310                3320                3330                3340
          *                   *                   *                   *                   *
TTT TCC TTA GGT GCT TCT CCA TAT CCT GGG GTA AAG ATT GAT GAA GAA TTT TGT
Phe Ser Leu Gly Ala Ser Pro Tyr Pro Gly Val Lys Ile Asp Glu Glu Phe Cys>

3350                3360                3370                3380                3390                3400
  *                   *                   *                   *                   *                   *
AGG CGA TTG AAA GAA GGA ACT AGA ATG AGG GCC CCT GAT TAT ACT ACA CCA GAA
Arg Arg Leu Lys Glu Gly Thr Arg Met Arg Ala Pro Asp Tyr Thr Thr Pro Glu>

            3410                3420                3430                3440                3450
              *                   *                   *                   *                   *
ATG TAC CAG ACC ATG CTG GAC TGC TGG CAC GGG GAG CCC AGT CAG AGA CCC ACG
Met Tyr Gln Thr Met Leu Asp Cys Trp His Gly Glu Pro Ser Gln Arg Pro Thr>

      3460                3470                3480                3490                3500                3510
        *                   *                   *                   *                   *                   *
TTT TCA GAG TTG GTG GAA CAT TTG GGA AAT CTC TTG CAA GCT AAT GCT CAG CAG
Phe Ser Glu Leu Val Glu His Leu Gly Asn Leu Leu Gln Ala Asn Ala Gln Gln>

              3520                3530                3540                3550                3560
                *                   *                   *                   *                   *
GAT GGC AAA GAC TAC ATT GTT CTT CCG ATA TCA GAG ACT TTG AGC ATG GAA GAG
Asp Gly Lys Asp Tyr Ile Val Leu Pro Ile Ser Glu Thr Leu Ser Met Glu Glu>
```

FIG. 7K

```
     3570           3580           3590           3600           3610
        *              *              *              *              *
GAT TCT GGA CTC TCT CTG CCT ACC TCA CCT GTT TCC TGT ATG GAG GAG GAG GAA
Asp Ser Gly Leu Ser Leu Pro Thr Ser Pro Val Ser Cys Met Glu Glu Glu Glu>

3620           3630           3640           3650           3660           3670
 *              *              *              *              *              *
GTA TGT GAC CCC AAA TTC CAT TAT GAC AAC ACA GCA GGA ATC AGT CAG TAT CTG
Val Cys Asp Pro Lys Phe His Tyr Asp Asn Thr Ala Gly Ile Ser Gln Tyr Leu>

         3680           3690           3700           3710           3720
            *              *              *              *              *
CAG AAC AGT AAG CGA AAG AGC CGG CCT GTG AGT GTA AAA ACA TTT GAA GAT ATC
Gln Asn Ser Lys Arg Lys Ser Arg Pro Val Ser Val Lys Thr Phe Glu Asp Ile>

   3730           3740           3750           3760           3770           3780
      *              *              *              *              *              *
CCG TTA GAA GAA CCA GAA GTA AAA GTA ATC CCA GAT GAC AAC CAG ACG GAC AGT
Pro Leu Glu Glu Pro Glu Val Lys Val Ile Pro Asp Asp Asn Gln Thr Asp Ser>

            3790           3800           3810           3820           3830
               *              *              *              *              *
GGT ATG GTT CTT GCC TCA GAA GAG CTG AAA ACT TTG GAA GAC AGA ACC AAA TTA
Gly Met Val Leu Ala Ser Glu Glu Leu Lys Thr Leu Glu Asp Arg Thr Lys Leu>

      3840           3850           3860           3870           3880
         *              *              *              *              *
TCT CCA TCT TTT GGT GGA ATG GTG CCC AGC AAA AGC AGG GAG TCT GTG GCA TCT
Ser Pro Ser Phe Gly Gly Met Val Pro Ser Lys Ser Arg Glu Ser Val Ala Ser>
```

19/28

FIG. 7L

```
3890       3900       3910       3920       3930       3940
  *          *          *          *          *          *
GAA GGC TCA AAC CAG ACA AGC GGC TAC CAG TCC GGA TAT CAC TCC GAT GAC ACA
Glu Gly Ser Asn Gln Thr Ser Gly Tyr Gln Ser Gly Tyr His Ser Asp Asp Thr>

        3950       3960       3970       3980       3990
          *          *          *          *          *
GAC ACC ACC GTG TAC TCC AGT GAG GAA GCA GAA CTT TTA AAG CTG ATA GAG ATT
Asp Thr Thr Val Tyr Ser Ser Glu Glu Ala Glu Leu Leu Lys Leu Ile Glu Ile>

 4000       4010       4020       4030       4040       4050
   *          *          *          *          *          *
GGA GTG CAA ACC GGT AGC ACA GCC CAG ATT CTC CAG CCT GAC ACG GGG ACC ACA
Gly Val Gln Thr Gly Ser Thr Ala Gln Ile Leu Gln Pro Asp Thr Gly Thr Thr>

       4060       4070
         *          *
CTG AGC TCT CCT CCT GTT TAA
Leu Ser Ser Pro Pro Val ***
```

FIG. 7M

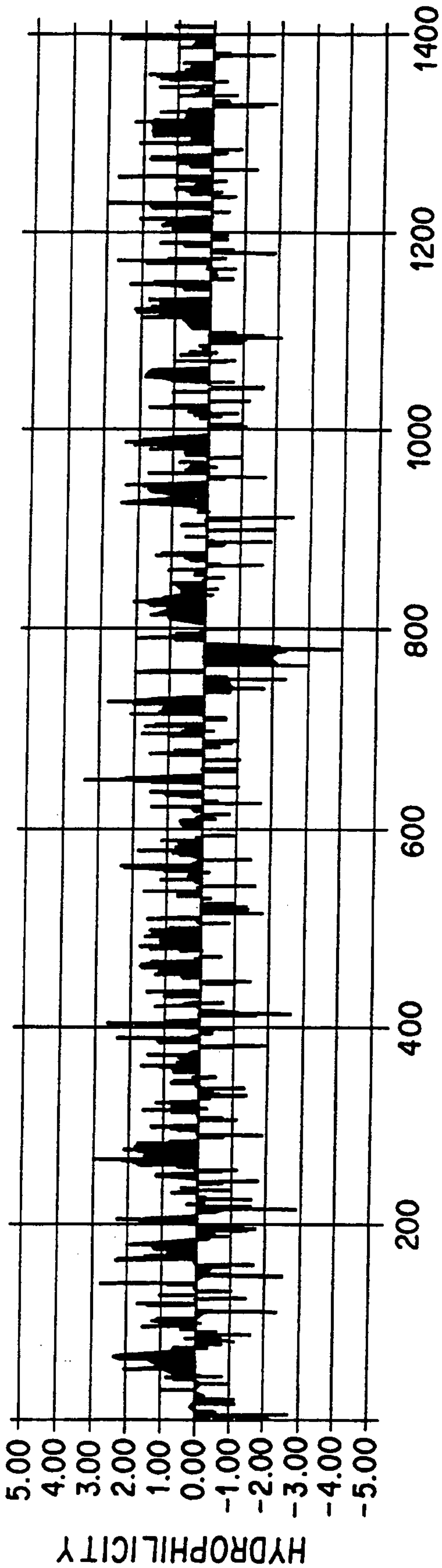

1400
1200
1000
800
600
400
200
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-5.00
HYDROPHILICITY

FIG. 8

```
                 * *   *
KDR   787  GTVKRANGGELKTGYLSIVMDPDELPLDEHCERLPYDASKWEFPRDRLNLGK
ckit  543  L***YLQKPMYEVQWKVVEEINGNNYVYIDPTQ****H-******N**SF**
CSF1  536  LLY*YKQKPKYQVRWKIIESYEGNSYTFIDPTQ***NE-******NN*QF**
PDGF  522  MLWQKKPRYEIRWKVIESVSSDGHEYIYVDPVQ****-ST******QLV**R

KDR   839  PLGRGAFGQEIEADAFGIDKTATCRTVAVKMLKEGATHSEHRALMSELKILI
ckit  594  T**A****KVVAET*Y*LI*SDAAM********PS*HLT*RE*******V*S
CSF1  587  T**A****KVV**T***LG*EDAVLK*******ST*HAD*KE********MS
PDGF  573  T**S*****VV**T*H*LSHSQATMK*******ST*RSS*KQS*******MS

KDR   891  HIGHHLNVVNLLGACTKPGGPLMVIVEFCKFGNLSTYLRSKRNEFVPYKTKG
ckit  646  YL*N*M*I********I-***TL**T*Y*CY*D*LNF**R**DS*ICS*QED
CSF1  639  *L*Q*E*I********H-***VL**T*Y*CY*D*LNF**R*AEAMLGPSLSP
PDGF  625  *L*P*************-***IYI*T*Y*RY*D*VD**HRNKHT*LQRHSNK

KDR   943  ARFRQGKDYVGAIPVDLKRRLDSIT-SSQSSASSGFVEEKSL-------SDV
ckit  697  HAEA-A-L*KNLLHSKESSCS-DS*N-E----YMDMKPGVS--YVVPT--KA
CSF1  690  GQDPE*GVDYKN*HLEK*YVRRDSGF***GVDTYVEMRPVSTSS-NDSF*EQ
PDGF  676  HCPPSAEL*SN*LP*GFSLPSHLNLTGESDGGYMDMSKDESIDYVPMLDMKG

KDR   987  EEEEAPEDLYKDF--------------------------LTLEHLICYSFQV
ckit  737  D-KRRSVRIGSYI-----------ERDVTPAIMEDDELA*D**D*LSF*Y**
CSF1  741  DLDKEDGRPL------------------------------E*RD*LHF*S**
PDGF  728  DIKY*DIESPSYMAPYDNYVPSAPERTYRATLINDSPV-*SYTD*VGF*Y**

KDR  1013  AKGMEFLASRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGD
```

FIG. 9A

```
ckit 777  -***A*****N*************THGRIT***********KN*SN**V**N
CSF1 762  *Q**A*****N**********V**TNGHVA**G********MN*SN*IV**N
PDGF 779  *N**D*****N**********V*ICEGKL************MR*SN*IS**S

KDR 1065  ARLPLKWMAPETIFDRVYTIQSDVWSFGVLLWEIFSLGASPYPGVKIDEEFC
ckit 828  ****V******S**NC***EE*****Y*IF********S*****MPVKSK*Y
CSF1 814  ****V******S***C***VQ*****Y*I*********LN****ILVNSK*Y
PDGF 831  ***********S**NS***TL*******I*********GT***ELPMNDQ*Y

KDR 1117  RRLKEGTRMRAPDYTTPEMYQTMLDCWHGEPSQRPTFSELVEHLGNLLQANA
ckit 880  KMI***F**LS*EHAPA***DI*KT**DAD*LK****KQIVQLIEKQISEST
CSF1 862  KLV*D*YQ*AQ*AFAPKNI*SI*QA**AL**TH****QQICSF*QEQAQEDR
PDGF 883  NAI*R*Y**AQ*AHASD*I*EI*QK**EEKFET**P**Q**LL*ER**GEGY

KDR 1169  QQDGKDYIVLPISETLSMEEDSGLSLPTSPVSCMEEEEVCDPKFHYDNTAGI
ckit 932  NHIYSNLANCSPNRQKPVVDHSVRINSVGSTASSSQPLLVHDDV
CSF1 914  RERDYTNLPSSSRSGG*GSSS*E*EEESSSEHLTCC*QGDIAQPLLQPNNYQ
PDGF 934  KKKYQQVDEEFLRSDHPAILR*QARF*GIHSLRSPLDTSSVLYTAVQPNESD

KDR 1213  SQYLQNSKRKSRPVSVKTFEDIPLEEPEVKVIPDDNQTDSGMVLASEELKTL
CSF1 966  FC
PDGF 987  ND*IIPLPDPKPD*ADEGLPEGSPSLASSTLNEVNTSSTISCDSPL*LQEEP

KDR 1273  EDRTKLSPSFGGMVPSKSRESVASEGSNQTSGYQSGYHSDDTDTTVYSSEEA
PDGF1039  QQAEPEAQLEQPQDSGCPGPLAEA*DSFLEQPQD**CPGPLAEAEDSFL

KDR 1325  ELLKLIEIGVQTGSTAQILQPDTGTTLSSPPV
```

FIG. 9B

IDENTIFICATION OF *kdp* mRNA

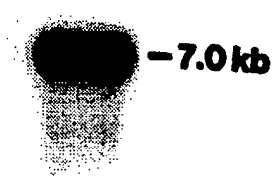
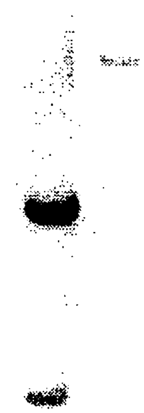


FIG.10

IDENTIFICATION OF *kdp* GENE
BY SOUTHERN ANALYSIS



1 2 3 4

FIG. 11

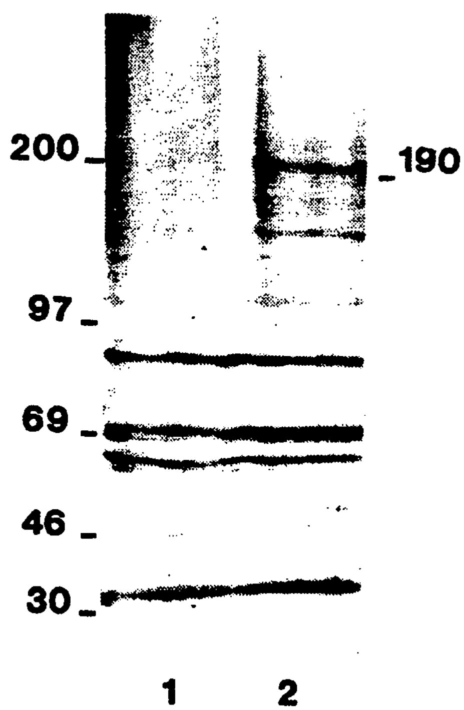


FIG. 12

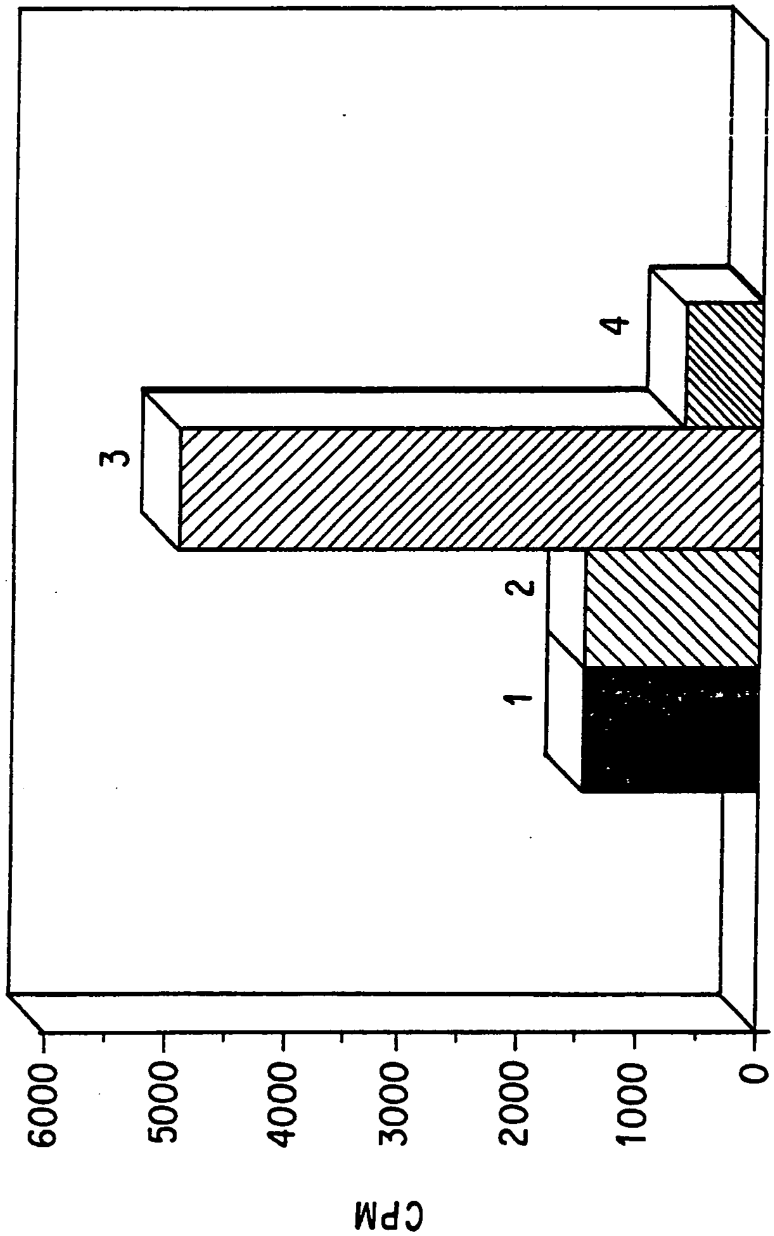


FIG. 13

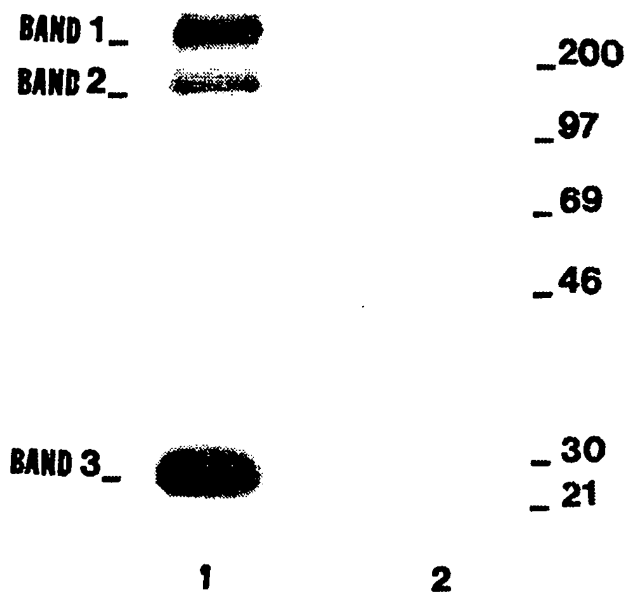


FIG. 14

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/01300

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C07K 3/00, 13/00; C07H 21/00; C12P 21/06, 21/02, 21/04; C12N 15/00
US CL : 530/387; 536/27; 435/69.1, 70.1, 71.1, 320.1

II. FIELDS SEARCHED

Minimum Documentation Searched[4]

Classification System	Classification Symbols
U.S.	530/387; 536/27; 435/69.1, 70.1, 71.1, 320.1

Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched[5]

APS, DIALOG
search terms: type III receptor tyrosine kinase

III. DOCUMENTS CONSIDERED TO BE RELEVANT[14]

Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages[17]	Relevant to Claim No.[18]
Y,P	Proc. Natl. Acad. Sci., Volume 88, Issued 1991, W. Mathewes et al., "A receptor tyrosine kinase cDNA isolated from a population of enriched primitive hematopoietic cells and exhibiting close genetic linkage to c-kit", pages 9026-9030, see entire document.	1-17
Y	Proc. Natl. Acad. Sci., Volume 86, Issued March 1989, A.F. Wilks, "Two putative protein-tyrosine kinases identified by applicatin of the polymerase chain reaction", pages 1603-1607, see entire document.	1-17
X,P	Oncogene, Volume 6, issued 1991, B.I. Terman et al., "Identification of a new endothelial cell growth factor receptor tyrosine kinase", pages 1677-1683, see entire document.	1-17
A	Oncogene, volume 3, issued 1988, M. Ruta et al., "A novel protein tyrosine kinase gene whose expression is modulated during endothelial cell differentiation", pages 9-15, see entire document.	1-17

* Special categories of cited documents:[15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search[2]	Date of Mailing of this International Search Report[2]
13 MAY 1992	19 MAY 1992
International Searching Authority[1]	**Signature of Authorized Officer**[20]
ISA/US	Lorraine M. Spector, Ph.D.

Form PCT/ISA/210 (second sheet)(May 1986) 8

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

A Y	Oncogene, volume 5, issued 1990, M. Shibuya et al., "Nucleotide sequence and expression of a novel human recept r-type tyrosine kinase gene (flt) clos ly related to the fms family", pag s 519-524, se entir document.	1-17 17
Y	Proc. Natl. Acad. Sci., Volume 85, Issued May 1988, R.G.K. Gronwald et al., "Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor: Evidence for more than one receptor class", pages 3435-3439, see entire document.	15

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE[1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers _, because they relate to subject matter (1) not required to be searched by this Authority, namely:

2. ☐ Claim numbers _, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out (1), specifically:

3. ☐ Claim numbers _, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING[2]

This International Searching Authority found multiple inventions in this international application as follows:

I. Claims 1-9 and 14-17, drawn to nucleic acids and expression thereof. Class 536, subclass 27 and Class 435, subclass 69.1.
II. Claims 10-13, drawn to an isolated growth factor receptor. Class 530, subclass 387.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application. (Telephone Practice)

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Search Authority did not invite payment of any additional fee.

Remark on protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No.[18]
Y	Proc. Natl. Acad. Sci., Volume 86, Issued November 1989, M. Streuli et al., "A family of receptor-linked protein tyrosine phosphatases in humans and Drosophila", pages 8698-8702, see entire document.	1-14
Y	M.A. Innes et al., PCR Protocols, a guide to methods and applications, published 1990 by Academic Press (N.Y.), see page 10.	15, 16